REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
03-05-2007	Biannual Technical Report	7/1/2007 - 12/31/2007

4. TITLE AND SUBTITLE

Bi-annual (7/1/2007--12/31/2007) Performance/Technical Report for ONR Award under Grant N00014-07-1-0395
MARSnet: Mission-aware Autonomous Radar Sensor Network for Future Combat Systems

5a. CONTRACT NUMBER

5b. GRANT NUMBER
N00014 – 07 –1 –0395

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Liang, Qilian

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
University of Texas at Arlington
Office of Sponsored Projects
PO Box 19145
Arlington, TX 76019

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Office of Naval Research
875 N. Randolph St.
One Liberty Center
Arlington, VA 22203-1995

10. SPONSOR/MONITOR'S ACRONYM(S)
ONR

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for Public Release; Distribution is Unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

During the period of 7/1/2007 -- 12/31/2007, we performed the following studies on radar sensor network: 1) Network-enabled Electronic Warfare (NEW) for Collaborative Automatic Target Recognition (CATR); 2) Foliage clutter modeling using narrowband and UWB radars; 3) A propagation Environment Modeling in Foliage using UWB radars; 4) Target detection in foliage using short-time Fourier transform and UWB radar sensor networks; 5) Some experimental studies on path loss models for wireless sensor networks based on Xbow motes, 6) Theoretical studies on distributed connected dominating set construction in random geometric k-Disk graphs for potential application to real sensor networks.

15. SUBJECT TERMS
Radar Sensor Network, UWB Radar, Sense through foliage, Automatic Target Recognition, clutter modeling.

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Qilian Liang
U	U	U	UU	85	19b. TELEPHONE NUMBER *(Include area code)* 817-272-1339

Bi-annual (7/1/2007–12/31/2007) Performance/Technical Report for ONR Award under Grant N00014-07-1-0395 MARSnet: Mission-aware Autonomous Radar Sensor Network for Future Combat Systems

Qilian Liang
Department of Electrical Engineering
University of Texas at Arlington
Arlington, TX 76019-0016 USA
Phone: 817-272-1339, Fax: 817-272-2253
E-mail: liang@uta.edu

Abstract

During the period of 7/1/2007 – 12/31/2007, we performed the following studies on radar sensor network:

1. Network-enabled Electronic Warfare (NEW) for Collaborative Automatic Target Recognition (CATR);
2. Foliage clutter modeling using narrowband and UWB radars;
3. A propagation Environment Modeling in Foliage using UWB radars;
4. Target detection in foliage using short-time Fourier transform and UWB radar sensor networks;
5. Some experimental studies on path loss models for wireless sensor networks based on Xbow motes,
6. Theoretical studies on distributed connected dominating set construction in random geometric k-Disk graphs for potential application to real sensor networks.

1 NEW-CATR: Network-enabled Electronic Warfare for Collaborative Automatic Target Recognition

Network-enabled Electronic Warfare (NEW) is to develop modeling and simulation effort to explore the advantages and limitations of network-enabled electronic warfare concepts. The advantages of linking multiple electronic support measures (ESM) and electronic attack (EA) assets to achieve improved capabilities across a networked battle force have yet to be quantified. In [1], we utilized radar sensors as ESM and EA assets to demonstrate the advantages of NEW in Collaborative Automatic Target Recognition (CATR). Signal (waveform) design for Radar Sensor Networks (RSN) in NEW is studied theoretically. The conditions for waveform coexistence and the interferences among waveforms in RSN are analyzed. We applied the NEW to CATR via waveform diversity combining and propose maximum-likelihood (ML)-ATR algorithms for non-fluctuating targets as well as fluctuating targets. Simulation results indicate that our NEW-CATR performs much better than the single sensor-based ATR algorithm for non-fluctuating targets and fluctuating targets.

Real world application example, sense-through-foliage target detection using UWB radars, was also used to validate our algorithms.

2 Foliage Clutter Modeling Using Narrowband and UWB Radars

In [2], we studied foliage clutter modeling on a basis of both pragmatically narrowband and ultra-wide band (UWB) radars. We proposed that the foliage clutter follows log-logistic model using maximum likelihood (ML) parameter estimation as well as the root mean square error (RMSE) on PDF curves between original clutter and statistical model data. In addition to investigating the log-logistic model, we also compared it with other popular clutter models, namely log-normal and Weibull. We showed that the log-logistic model not only achieves the smallest standard deviation (STD) error on estimated model parameters, but also has the best goodness-of-fit and smallest RMSE.

3 A Propagation Environment Modeling in Foliage Using UWB Radar

In [3], we proposed that foliage clutter follows log-logistic model using maximum likelihood (ML) parameter estimation as well as the root mean square error (RMSE) on PDF curves between original clutter and statistical model data. We not only investigated log-logistic model, but compared it with other popular clutter models, namely log-normal, Weibull and Nakagami. It showed that the log-logistic model not only achieves the small- est standard divination (STD) error on estimated model parameters, but also the best goodness-of-fit and smallest RMSE for both poor and good clutter signals.

4 Target Detection in Foliage Using Short-Time Fourier Transform and UWB Radar Sensor Networks

In [4], we studied sense-through-foliage target detection. When radar echoes are in good quality, the detection of target can be achieved by applying short time Fourier transform (STFT) to the received UWB radar waveform. We compared our approach in case of no target as well as with target against the scheme in which 2-D image was created via adding voltages with the appropriate time offset. Results show that our approach can detect target more easily. When radar echoes are in poor condition and single radar is unable to carry out the detection, we employed both Radar Sensor Networks (RSN) and RAKE structure to combine the echoes from different radar members and finally detected the target.

5 Experimental Path Loss Models for Wireless Sensor Networks

Energy conservation is critical in Wireless Sensor Networks. Replacing or recharging batteries is not an option for sensors deployed in hostile environments. Generally communication electronics in the sensor utilizes most energy. In [5], we studied the effect of changing the transmission power and baud rate on transmission distance. Using Shannon channel capacity formula and LogDistance Path Loss Model, transmission distance is shown to be related to transmit power and baud rate. Extensive empirical readings were taken to confirm the above relation. The path loss exponent got as a result of data fitting is within the acceptable range for wireless environment. Using the equation

derived in [5], the distance between neighboring motes and traffic density, it will be possible for sensors to adjust their transmit power and baud rate so as to use only the required amount of energy to maintain the wireless link to the neighbor and conserve power.

6 Distributed Connected Dominating Set Construction in Random Geometric k-Disk Graphs

In [6], we studied the problem of minimum connected dominating set in random geometric k-disk graphs. This research is motivated by the problem of virtual backbone construction in wireless ad hoc and sensor networks, where the coverage area of nodes are disks with different radii. We derived the size relationship of any maximal independent set and minimum connected dominating set in geometric k-disk graphs, and applied it to analyze the performances of two distributed connected dominating set algorithms we proposed in [6]. These algorithms have bounded performance ratio and low communication overhead, and therefore have the potential to be applied in real ad hoc and sensor networks

References

[1] Q. Liang, X. Cheng, S. Samn, "NEW-CATR: Network-enabled Electronic Warfare for Collaborative Automatic Target Recognition," submitted to *IEEE Trans on Aerospace and Electronic Systems.*

[2] J. Liang, Q. Liang, "Foliage Clutter Modeling Using Narrowband and UWB Radars," submitted to *IEEE Globecom*, New Orleans, LA, 2008.

[3] J. Liang, Q. Liang, S. Samn, "A Propagation Environment Modeling in Foliage," submitted to *IEEE Trans on Antenna and Propagation.*

[4] J. Liang, Q. Liang, "Target Detection in Foliage Using Short-Time Fourier Transform and UWB Radar Sensor Networks," submitted to *IEEE Globecom*, New Orleans, LA, 2008.

[5] R. Sawant, Q. Liang, D. Popa, and F. Lewis, "Experimental Path Loss Models for Wireless Sensor Networks," presented at *IEEE MILCOM*, Orlando, FL, Oct 2007.

[6] D. Chen, K. Xing, Q. Liang, and E. Park, "Distributed Connected Dominating Set Construction in Random Geometric k-Disk Graphs," accepted by *IEEE International Conference on Distributed Computing Systems (ICDCS)*, Beijing, China, June 2008.

NEW-CATR: Network-enabled Electronic Warfare for Collaborative Automatic Target Recognition

Qilian Liang
Dept of Electrical Engineering
University of Texas at Arlington
Arlington, TX 76019-0016, USA
E-mail: liang@uta.edu

Xiuzhen Cheng
Dept of Computer Science
George Washington University
Washington, DC 20052, USA
E-mail: cheng@gwu.edu

Sherwood W. Samn
Air Force Research Laboratory/HEX
Brooks City Base
San Antonio, TX 78235, USA
E-mail: Sherwood.samn@brooks.af.mil

Abstract

Network-enabled Electronic Warfare (NEW) is to develop modeling and simulation effort to explore the advantages and limitations of network-enabled electronic warfare concepts. The advantages of linking multiple electronic support measures (ESM) and electronic attack (EA) assets to achieve improved capabilities across a networked battle force have yet to be quantified. In this paper, we utilize radar sensors as ESM and EA assets to demonstrate the advantages of NEW in Collaborative Automatic Target Recognition (CATR). Signal (waveform) design for Radar Sensor Networks (RSN) in NEW is studied theoretically. The conditions for waveform coexistence and the interferences among waveforms in RSN are analyzed. We apply the NEW to CATR via waveform diversity combining and propose maximum-likelihood (ML)-ATR algorithms for non-fluctuating targets as well as fluctuating targets. Simulation results indicate that our NEW-CATR performs much better than the single sensor-based ATR algorithm for non-fluctuating targets and fluctuating targets.

Index Terms : Network-enabled electronic warfare, radar sensor networks, waveform diversity, collaborative automatic target recognition, maximum-likelihood, interferences.

1 Introduction and Motivation

In current and future military operational environments such as Global War on Terrorism (GWOT) and Maritime Domain Awareness (MDA), war fighters require technology that can support their information needs in manner that is independent of their location and consistent with their level of command or responsibility and operational situation. To support this need, the U.S. Department of Defense (DoD) has developed the concept of Network Centric Warfare (NCW), defined as "*military operations that exploit state-of-the-art information and networking technology to integrate widely dispersed human decision makers, situational and targeting sensors, and forces and weapons into a highly adaptive, comprehensive system to achieve unprecedented mission effectiveness*" [1]. The goal of electronic warfare is to control the electromagnetic (EM) spectrum by disrupting or denying enemy use of the spectrum while ensuring its use by friendly forces [2].

Network-enabled Electronic Warfare (NEW) is to develop modeling and simulation effort to explore the advantages and limitations of network-enabled electronic warfare concepts. The advantages of linking multiple electronic support measures (ESM) and electronic attack (EA) assets to achieve improved capabilities across a networked battle force have yet to be quantified [2]. In this paper, we utilize radar sensors as ESM and EA assets to demonstrate the advantages of NEW in Collaborative Automatic Target Recognition (CATR). The network of radar sensors should operate with multiple goals managed by an intelligent platform network that can manage the dynamics of each radar to meet the common goals of the platform, rather than each radar operates as an independent system. Therefore, it is significant to perform signal design and processing and networking cooperatively within and between platforms of radar sensors and their communication modules. This need is also testified by the recent solicitation from the U.S. Office of Naval Research [2][3]. For example, in [3], it is stated that "Algorithms are sought for fused, and/or, coherent cross-platform Radio Frequency (RF) sensing. The focus of this effort is to improve surveillance utilizing

a network, not fusion of disparate sensor products. The algorithms should be capable of utilizing RF returns from multiple aspects in a time-coordinated sensor network." In this paper, we study waveform design and diversity algorithms for radar sensor networks. Waveform diversity is the technology that allows one or more sensors on board a platform to automatically change operating parameters, e.g., frequency, gain pattern, and pulse repetition frequency (PRF), to meet the varying environments. It has long been recognized that judicious use of properly designed waveforms, coupled with advanced receiver strategies, is fundamental to fully utilizing the capacity of the electromagnetic spectrum. However, it is the relatively recent advances in hardware technology that are enabling a much wider range of design freedoms to be explored. As a result, there are emerging and compelling changes in system requirements such as more efficient spectrum usage, higher sensitivities, greater information content, improved robustness to errors, reduced interference emissions, etc. The combination of these changes is fueling a worldwide interest in the subject of waveform design and the use of waveform diversity techniques.

In the existing works on waveform design and selection, Fitzgerald [8] demonstrated the inappropriateness of selection of waveforms based on measurement quality alone: the interaction between the measurement and the track can be indirect, but must be accounted for. Bell [6] used information theory to design radar waveforms for the measurement of extended radar targets exhibiting resonance phenomena. In [5], the singularity expansion method was used to design discriminant waveforms such as K-pulse, E-pulse, and S-pulse. Sowelam and Tewfik [24] developed a signal selection strategy for radar target classification, and a sequential classification procedure was proposed to minimize the average number of necessary signal transmissions. Intelligent waveform selection was studied in [4][12], but the effect of Doppler shift was not considered. In [16], time-frequency-based generalized chirps were used as waveform for detection and estimation. In [15], the performance of constant frequency (CF) and linear frequency modulated (LFM) waveform fusion

from the standpoint of the whole system was studied, but the effect of clutter was not considered. In [23], CF and LFM waveforms were studied for a sonar system, but it was assumed that the sensor is non-intelligent (i.e., waveform can't be selected adaptively). All the above studies and design methods focused on the waveform design or selection for a single active radar or sensor. In [21], cross-correlation properties of two radars were briefly mentioned and the binary coded pulses using simulated annealing [7] are highlighted. However, the cross-correlation of two binary sequences such as binary coded pulses (e.g. Barker sequence) is much easier to study than that of two analog radar waveforms. In this paper, we focus on the waveform diversity and design for radar sensor networks using the constant frequency (CF) pulse waveform. Relative to previous work, this paper has the following novelties:

1. A networked scenario of radar sensors is considered instead of a single radar system.

2. We study waveform design and diversity for radar sensors networks. In Space-Time Adaptive Processing (STAP) [18], the waveform (pulse) design is essentially for a single radar sensor. The pulse is sent repeatedly at different time and echo is received and processed by the antenna array, and no interferences among pulses if the pulse repetition interval is large enough.

3. We investigate collaborative automatic target recognition using radar sensor networks and compare it against single radar system in CATR.

4. Simulations are performed for nonfluctuating targets as well as fluctuating targets, and a real world application example, sense-through-foliage target detection, is presented.

The rest of this paper is organized as follows. In Section 2, we study the coexistence of radar waveforms. In Section 3, we analyze the interferences among radar waveforms. In Section 4, we propose a RAKE structure for waveform diversity combining and present maximum-likelihood (ML) algorithms for CATR. In Section 5, we provide simulation results on ML-CATR. In Section 6, we

conclude this paper and discuss future research.

2 Co-existence of Radar Waveforms

In radar sensor networks (RSN), radar sensors interfere with each other and the signal-to-interference-ratio may be very low if the waveforms are not properly designed. In this paper, we introduce orthogonality as one criterion for waveform design in RSN to make radars coexistence. In addition, since the radar channel is narrow-band, we will also consider the bandwidth constraint.

In our radar sensor networks, we choose the CF pulse waveform, which can be defined as

$$x(t) = \sqrt{\frac{E}{T}} \exp(j2\pi\beta t) \quad -T/2 \leq t \leq T/2 \tag{1}$$

where β is the RF carrier frequency in radians per second. In radar, ambiguity function (AF) is an analytical tool for waveform design and analysis, which succinctly characterizes the behavior of a waveform paired with its matched filter. The ambiguity function is useful for examining resolution, side lobe behavior, and ambiguities in both range and Doppler for a given waveform [18]. For a single radar, the matched filter for waveform $x(t)$ is $x^*(-t)$, and the ambiguity function of CF pulse waveform is

$$\begin{aligned} A(\tau, F_D) &= \left| \int_{-T/2+\tau}^{T/2} x(t) \exp\left(j2\pi F_D s\right) x^*(t-\tau) dt \right| \\ &= \left| \frac{E \sin[\pi F_D (T - |\tau|)]}{T\pi F_D} \right| \quad -T \leq \tau \leq T \end{aligned} \tag{2}$$

We can simplify this AF in the following three special cases:

1. When $\tau = 0$,

$$A(0, F_D) = \left| \frac{E \sin(\pi F_D T)}{T\pi(F_D)} \right|; \tag{3}$$

2. when $F_D = 0$,

$$A(\tau, 0) = \left| \frac{E(T - |\tau|)}{T} \right|; \tag{4}$$

3. and when $\tau = F_D = 0$,

$$A(0,0) = E \tag{5}$$

Note that the above ambiguity is for one radar only (no coexisting radar).

For radar sensor networks, the waveforms from different radars interfere with each other. We choose the waveform for radar i as

$$x_i(t) = \sqrt{\frac{E}{T}} \exp[j2\pi(\beta + \delta_i)t] \quad -T/2 \leq t \leq T/2 \tag{6}$$

which means that there is a frequency shift δ_i for radar i. To minimize the interference from one waveform to another, optimal values for δ_i should be determined to make the waveforms orthogonal to each other, i.e., let the cross-correlation between $x_i(t)$ and $x_n(t)$ be 0,

$$\int_{-T/2}^{T/2} x_i(t)x_n^*(t)dt = \frac{E}{T}\int_{-T/2}^{T/2} \exp[j2\pi(\beta + \delta_i)t] \exp[-j2\pi(\beta + \delta_n)t]dt \tag{7}$$

$$= E\text{sinc}[\pi(\delta_i - \delta_n)T] \tag{8}$$

If we choose

$$\delta_i = \frac{i}{T} \tag{9}$$

where i is a dummy index, (8) can be written as

$$\int_{-T/2}^{T/2} x_i(t)x_n^*(t)dt = \begin{cases} E & i = n \\ 0 & i \neq n \end{cases} \tag{10}$$

Therefore choosing $\delta_i = \frac{i}{T}$ in (6) yields orthogonal waveforms, i.e., the waveforms can coexist if the carrier spacing is a multiple of $1/T$ between two radar waveforms. In other words, orthogonality amongst carriers can be achieved by separating the carriers by a multiple of the inverse of waveform pulse duration. With this design, all the orthogonal waveforms can work simultaneously. However, there may exist time delay and Doppler shift ambiguity which may interfere with other waveforms in RSN.

3 Interferences of Waveforms In Radar Sensor Networks

3.1 RSN with Two Radar Sensors

We are interested in analyzing the interference from one radar to another if there exist time delay and Doppler shift. For a simple case where there are two radar sensors (i and n), the ambiguity function of radar i (considering the interference from radar n) is

$$A_i(t_i, t_n, F_{D_i}, F_{D_n}) = \left| \int_{-\infty}^{\infty} [x_i(t) \exp(j2\pi F_{D_i} t) + x_n(t - t_n) \exp(j2\pi F_{D_n} t)] x_i^*(t - t_i) dt \right| \quad (11)$$

$$\leq \left| \int_{-T/2+\max(t_i,t_n)}^{T/2+\min(t_i,t_n)} x_n(t - t_n) \exp(j2\pi F_{D_n} t) x_i^*(t - t_i) dt \right| + \left| \int_{-T/2+t_i}^{T/2} x_i(t) \exp(j2\pi F_{D_i} t) x_i^*(t - t_i) dt \right| \quad (12)$$

$$= \left| \int_{-T/2+\max(t_i,t_n)}^{T/2+\min(t_i,t_n)} x_n(t - t_n) \exp(j2\pi F_{D_n} t) x_i^*(t - t_i) dt \right| + \left| \frac{E \sin[\pi F_{D_i}(T - |t_i|)]}{T\pi F_{D_i}} \right| \quad (13)$$

To make the analysis easier, it is generally assumed that the radar sensor platform has access to the Global Positioning Service (GPS) and the Inertial Navigation Unit (INU) timing and navigation data [3]. In this paper, we assume that the radar sensors are synchronized and that $t_i = t_n = \tau$. Then (13) can be simplified as

$$A_i(\tau, F_{D_i}, F_{D_n}) \approx |E\text{sinc}[\pi(n - i + F_{D_n} T)]| + \left| \frac{E \sin[\pi F_{D_i}(T - |\tau|)]}{T\pi F_{D_i}} \right| \quad (14)$$

We have the following three special cases:

1. If $F_{D_i} = F_{D_n} = 0$, and δ_i and δ_n follow (9), (14) becomes

$$A_i(\tau, 0, 0) \approx \left| \frac{E(T - |\tau|)}{T} \right| \quad (15)$$

2. If $\tau = 0$, (14) becomes

$$A_i(0, F_{D_i}, F_{D_n}) \approx |E\text{sinc}[\pi(n - i + F_{D_n} T)]| + \left| \frac{E \sin(\pi F_{D_i} T)}{T\pi F_{D_i}} \right| \quad (16)$$

3. If $F_{D_i} = F_{D_n} = 0$, $\tau = 0$, and δ_i and δ_n follow (9), (14) becomes

$$A_i(0,0,0) \approx E \tag{17}$$

3.2 RSN with M Radar Sensors

Our analysis on an RSN with two radar sensors can be extended to the case of M radars. Assuming that the time delay τ for each radar is the same, then the ambiguity function of radar 1 (considering interferences from all the other $M-1$ radars with CF pulse waveforms) can be expressed as

$$A_1(\tau, F_{D_1}, \cdots, F_{D_M}) \approx \sum_{i=2}^{M} |E\text{sinc}[\pi(i-1+F_{D_i}T)]| + \left| \frac{E\sin[\pi F_{D_1}(T-|\tau|)]}{T\pi F_{D_1}} \right| \tag{18}$$

Similarly, we have the following three special cases:

1. $F_{D_1} = F_{D_2} = \cdots = F_{D_M} = 0$, and the frequency shift δ_i in (6) for each radar follows (9), then (18) becomes

$$A_1(\tau, 0, 0, \cdots, 0) \approx \left| \frac{E(T-|\tau|)]}{T} \right| \tag{19}$$

   Comparing it against (4), we notice that a radar may exist that can get the same signal strength as that of the single radar in a single radar system (no coexisting radar) when the Doppler shift is 0.

2. If $\tau = 0$, then (18) becomes

$$A_1(0, F_{D_1}, F_{D_2}, \cdots, F_{D_M}) \approx \sum_{i=2}^{M} |E\text{sinc}[\pi(i-1+F_{D_i}T)]| + \left| \frac{E\sin(\pi F_{D_1}T)}{T\pi F_{D_1}} \right| \tag{20}$$

   Comparing to (3), a radar in RSN has higher interferences when unknown Doppler shifts exist.

3. $F_{D_1} = F_{D_2} = \cdots = F_{D_M} = 0$, $\tau = 0$, and δ_i in (6) follows (9), then (18) becomes

$$A_1(0, 0, 0, \cdots, 0) \approx E \tag{21}$$

4 NEW for Collaborative Automatic Target Recognition

In NEW, the radar sensors are networked together in an ad hoc fashion. They do not rely on a preexisting fixed infrastructure, such as a wireless backbone network or a base station. They are self-organizing entities that are deployed on demand in support of various events surveillance, battlefield, disaster relief, search and rescue, etc. Scalability concern suggests a hierarchical organization of radar sensor networks with the lowest level in the hierarchy being a cluster. As argued in [9] [10] [13] [17], in addition to helping with scalability and robustness, aggregating sensor nodes into clusters has additional benefits:

1. conserving radio resources such as bandwidth;
2. promoting spatial code reuse and frequency reuse;
3. simplifying the topology, e.g., when a mobile radar changes its location, it is sufficient for the nodes in the attended clusters to update their topology information;
4. reducing the generation and propagation of routing information; and,
5. concealing the details of global network topology from individual nodes.

In RSN, each radar can provide its waveform parameters such as δ_i to its clusterhead radar, and the clusterhead radar can combine the waveforms from its cluster members.

In RSN with M radars, the received signal for clusterhead (assume it's radar 1) is

$$r_1(u,t) = \sum_{i=1}^{M} \alpha(u) x_i(t - t_i) \exp(j2\pi F_{D_i} t) + n(u,t) \tag{22}$$

where $\alpha(u)$ stands for radar cross section (RCS), which can be modeled using non-zero constants for non-fluctuating targets and four Swerling target models for fluctuating targets [18]; F_{D_i} is the Doppler shift of the target relative to waveform i; t_i is the delay of waveform i, and $n(u,t)$ is the

additive white Gaussian noise (AWGN). In this paper, we propose a RAKE structure for waveform diversity combining, as illustrated by Fig. 1. This figure summarizes how the clusterhead works. The received signal $r_1(u,t)$ consists of echos triggered by the waveforms from each radar sensor, and $x_i^*(t-t_i)$ is used to retrieve the amplified waveform from radar i (amplified by the target RCS) based on the orthogonal property presented in Sections 2 and 3, and then this information is time-averaged for diversity combining.

According to this structure, the received $r_1(u,t)$ is processed by a bank of matched filters, then the output of branch 1 (after integration) is

$$\begin{aligned} & Z_1(u;t_1,\cdots,t_M,F_{D_1},\cdots,F_{D_M}) \\ = & \int_{-T/2}^{T/2} r_1(u,t)x_1^*(t-t_1)ds \qquad (23) \\ = & \int_{-T/2}^{T/2}[\sum_{i=1}^{M}\alpha_i(u)x_i(t-t_i)\exp(j2\pi F_{D_i}t)+n(u,t)]x_1^*(t-t_1)dt \qquad (24) \end{aligned}$$

Assuming $t_1=t_2=\cdots=t_M=\tau$, then based on (18),

$$\begin{aligned} Z_1(u;\tau,F_{D_1},\cdots,F_{D_M}) \approx & \sum_{i=2}^{M}\alpha(u)E\text{sinc}[\pi(i-1+F_{D_i}T)] \\ + & \frac{\alpha(u)E\sin[\pi F_{D_1}(T-|\tau|)]}{T\pi F_{D_1}}+n(u,\tau) \qquad (25) \end{aligned}$$

Similarly, we can get the output for any branch m ($m=1,2,\cdots,M$),

$$\begin{aligned} Z_m(u;\tau,F_{D_1},\cdots,F_{D_M}) \approx & \sum_{i=1,i\neq m}^{M}\alpha(u)E\text{sinc}[\pi(i-m+F_{D_i}T)] \\ + & \frac{\alpha(u)E\sin[\pi F_{D_m}(T-|\tau|)]}{T\pi F_{D_m}}+n(u,\tau) \qquad (26) \end{aligned}$$

Therefore $Z_m(u;\tau,F_{D_1},\cdots,F_{D_M})$ consists of three parts, namely signal (reflected signal from radar m waveform): $\frac{\alpha(u)E\sin[\pi F_{D_m}(T-|\tau|)]}{T\pi F_{D_m}}$, interferences from other waveforms: $\sum_{i=1,i\neq m}^{M}\alpha(u)E\text{sinc}[\pi(i-m+F_{D_i}T)]$, and noise: $n(u,\tau)$.

We can also have the following three special cases for $|Z_m(u;\tau,F_{D_1},\cdots,F_{D_M})|$:

1. When $F_{D_1} = \cdots = F_{D_M} = 0$,

$$Z_m(u;\tau,0,0,\cdots,0) \approx \frac{E\alpha(u)(T-|\tau|)]}{T} + n(u,\tau) \tag{27}$$

which means that if there is no Doppler mismatch, there is no interference from other waveforms.

2. If $\tau = 0$, (26) becomes

$$\begin{aligned} & Z_m(u;0,F_{D_1},\cdots,F_{D_M}) \\ \approx & \sum_{i=1,i\neq m}^{M} \alpha(u)E\text{sinc}[\pi(i-m+F_{D_i}T)] + \frac{\alpha(u)E\sin[\pi F_{D_m}T]}{T\pi F_{D_m}} + n(u) \end{aligned} \tag{28}$$

3. If $\tau = 0$, and $F_{D_1} = \cdots = F_{D_M} = 0$, (26) becomes

$$Z_m(u;0,0,0,\cdots,0) \approx E\alpha(u) + n(u) \tag{29}$$

Doppler mismatch happens quite often in target search where target velocity is not yet known. However, in target recognition, generally high-resolution measurements of targets in range ($\tau = 0$) and Doppler are available, therefore (29) will be used for CATR.

How to combine all the Z_m's ($m = 1,2,\cdots,M$) is very similar to the diversity combining in wireless communications to combat channel fading, and the combination schemes may be different for different applications. In this paper, we are interested in applying the RSN waveform diversity to CATR, e.g., recognition that the echo on a radar display is that of an aircraft, ship, motor vehicle, bird, person, rain, chaff, clear-air turbulence, land clutter, sea clutter, bare mountains, forested areas, meteors, aurora, ionized media, or other natural phenomena via collaborations among different radars. Early radars were "blob" detectors in that they detected the presence of a target and gave its location in range and angle, and radar began to be more than a blob detector and could provide recognition of one type of target from another [21]. It is known that small changes in

the aspect angle of complex (multiple scatter) targets can cause major changes in the radar cross section (RCS). This has been considered in the past as a means of target recognition, and is called *fluctuation of radar cross section with aspect angle*, but it has not had much success [21]. In [19], a parametric filtering approach was proposed for target detection using airborne radar. In [14], knowledge-based sensor networks were applied to threat assessment. In this paper, we propose maximum likelihood collaborative automatic target recognition (ML-CATR) algorithms for RSN. We study non-fluctuating targets as well as fluctuating targets.

4.1 ML-CATR for Non-fluctuating Targets

In some sources, the non-fluctuating target is identified as a "Swerling 0" or "Swerling 5" model [22]. For non-fluctuating targets, the RCS $\alpha_m(u)$ is just a constant α for a given target. In (29), $n(u,\tau)$ is a zero-mean Gaussian random variable for a given τ, so $|Z_m(u;0,0,\cdots,0)|$ follows a Rician distribution because signal $E\alpha(u)$ is a positive constant $E\alpha$ for non-fluctuating targets. Let $y_m \stackrel{\triangle}{=} |Z_m(u;0,0,\cdots,0)|$, then the probability density function (pdf) of y_m is

$$f(y_m) = \frac{2y_m}{\sigma^2}\exp[-\frac{(y_m^2+\lambda^2)}{\sigma^2}]I_0(\frac{2\lambda y_m}{\sigma^2}) \tag{30}$$

where

$$\lambda = E\alpha, \tag{31}$$

σ^2 is the noise power (with I and Q sub-channel power $\sigma^2/2$), and $I_0(\cdot)$ is the zero-order modified Bessel function of the first kind. Let $\mathbf{y} \stackrel{\triangle}{=} [y_1, y_2, \cdots, y_M]$, then the pdf of $\mathbf{y}$ is

$$f(\mathbf{y}) = \prod_{m=1}^{M} f(y_m) \tag{32}$$

Our CATR is a multiple-category hypothesis testing problem, i.e., to decide a target category (e.g. aircraft, ship, motor vehicle, bird, etc.) based on $r_1(u,t)$. Assume there are totally N categories and a category n target has RCS α_n, therefore the ML-CATR algorithm to decide a

target category C can be expressed as,

$$C = \arg\max_{n=1}^{N} f(\mathbf{y}|\lambda = E\alpha_n) \tag{33}$$

$$= \arg\max_{n=1}^{N} \prod_{m=1}^{M} \frac{2y_m}{\sigma^2} \exp[-\frac{(y_m^2 + E^2\alpha_n^2)}{\sigma^2}] I_0(\frac{2E\alpha_n y_m}{\sigma^2}) \tag{34}$$

4.2 ML-CATR for Fluctuating Targets

Fluctuating target modeling is more realistic in which the target RCS is drawn from either the Rayleigh or chi-square of degree four pdf. The Rayleigh model describes the behavior of a complex target consisting of many scatters, none of which is dominant. The fourth-degree chi-square models targets having many scatters of similar strength with one dominant scatter. Based on different combinations of pdf and decorrelation characteristics (scan-to-scan or pulse-to-pulse decorrelation), four Swerling models are used [18]. In this paper, we focus on the "Swerling 2" model which is a Rayleigh distribution with pulse-to-pulse decorrelation. The pulse-to-pulse decorrelation implies that each individual pulse results in an independent value for RCS α.

For the Swerling 2 model, the RCS $|\alpha(u)|$ follows a Rayleigh distribution and its I and Q subchannels follow zero-mean Gaussian distributions with a variance γ^2. Assume

$$\alpha(u) = \alpha_I(u) + j\alpha_Q(u) \tag{35}$$

and $n(u) = n_I(u) + jn_Q(u)$ follows a zero-mean complex Gaussian distribution with a variance σ^2 for the I and Q subchannels. Therefore according to (29), $Z_m(u; 0, 0, 0, \cdots, 0)$ is a zero-mean Gaussian random variable with a variance $E^2\gamma^2 + \sigma^2$ for the I and Q subchannels, which means $y_m \triangleq |Z_m(u; 0, 0, \cdots, 0)|$ follows a Rayleigh distribution with a parameter $\sqrt{E^2\gamma^2 + \sigma^2}$,

$$f(y_m) = \frac{y_m}{E^2\gamma^2 + \sigma^2} \exp(-\frac{y_m^2}{E^2\gamma^2 + \sigma^2}) \tag{36}$$

The mean value of y_m is $\sqrt{\frac{\pi(E^2\gamma^2+\sigma^2)}{2}}$, and the variance is $\frac{(4-\pi)(E^2\gamma^2+\sigma^2)}{2}$. The variance of signal is $\frac{(4-\pi)E^2\gamma^2}{2}$ and the variance of noise is $\frac{(4-\pi)\sigma^2}{2}$.

Let $\mathbf{y} \stackrel{\triangle}{=} [y_1, y_2, \cdots, y_M]$, then the pdf of $\mathbf{y}$ is

$$f(\mathbf{y}) = \prod_{m=1}^{M} f(y_m) \tag{37}$$

Assume there are totally N categories and a category n target has a RCS $\alpha_n(u)$ (with a variance γ_n^2), so the ML-ATR algorithm to decide a target category C can be expressed as,

$$C = \arg\max_{n=1}^{N} f(\mathbf{y}|\gamma = \gamma_n) \tag{38}$$

$$= \arg\max_{n=1}^{N} \prod_{m=1}^{M} \frac{y_m}{E^2\gamma_n^2 + \sigma^2} \exp(-\frac{y_m^2}{E^2\gamma_n^2 + \sigma^2}) \tag{39}$$

5 Simulations and Real World Application Example

5.1 Computer Simulations

Radar sensor networks will be required to detect a broad range of target classes. Too often, the characteristics of objects that are not of interest (e.g., bird) are similar to those of threat objects (e.g., missile). Therefore, new techniques to discriminate threat against undesired detections (e.g. birds, etc.) are needed. We applied our ML-CATR to this important application, to recognize a target from many target classes. We assume that the domain of target classes is known a priori (N in Sections 4.1 and 4.2), and that the RSN is confined to work only on the known domain.

For non-fluctuating target recognition, our targets have 5 classes with different RCS values, which are summarized in Table 1 [21]. We applied the ML-CATR algorithms in Section 4.1 (for the non-fluctuating target case) to classify an unknown target as one of these 5 target classes. At each average SNR value, we ran Monte-Carlo simulations for 10^5 times for each target. The average SNR value is based on the average power from all targets (signal variance), so the actual SNRs for bird and missile are much lower than the average SNR value. For example, at the average SNR=16dB, the bird target SNR=-33.1646dB, and the missile target SNR=0.8149dB; and

at average SNR=20dB, the bird target SNR=-29.1646dB, and the missile target SNR=4.8149dB. In Fig. 2(a)(b), we plotted the probability of the ATR error in bird and missile recognition when they are assumed as non-fluctuating targets. These figures indicate that a single radar system can't perform well in both recognitions, whose probability of the ATR error is above 10%, which can't be used for real-world ATR. However, the 5-radar RSN and 10-radar RSN can maintain very low ATR errors. In Fig. 2(c), we plotted the average probability of the ATR error for all 5 targets recognition. Since the other 3 targets (different aircrafts) have much higher SNRs, their ATR error is lower, which makes the average probability of ATR error lower.

For fluctuating target recognition, we assume the fluctuating targets follow the "Swerling 2" model (Rayleigh distribution with pulse-to-pulse decorrelation), and assume the RCS value listed in Table 1 to be the standard deviation (std) γ_n of RCS $\alpha_n(u)$ for target n. We applied the ML-CATR algorithm in Section 4.2 (for the fluctuating target case) for target recognition within the 5 targets domain. Similarly we ran Monte-Carlo simulations at each SNR value. In Fig. 3(a)(b)(c), we plotted the ATR performance for fluctuating targets and compared the performances of a single-radar system, a 5-radar RSN, and a 10-radar RSN. Observe that the two RSNs perform much better than the single radar system. The ATR error for the missile is higher than that of bird because the Rayleigh distribution of the missile has a lot of overlap with its neighbor targets (aircrafts). Comparing Fig. 2(a)(b)(c) to Fig. 3(a)(b)(c), it is clear that higher SNRs are needed for the fluctuating target recognition comparing to the non-fluctuating target recognition. According to Skolnik [21], the radar performance with a probability of recognition error (p_e) less than 10% is good enough. Our RSN with waveform-diversity can achieve a probability of ATR error much less than 10% for each target ATR as well as the average ATR for all targets. However, the single radar system has a probability of ATR error much higher than 10%. Fig. 3(c) also tells us that the average probability of ATR error of a single-radar system is impossible to be less than 10% even at

an extremely high SNR. Our RSN with waveform diversity is very promising for real-world ATR.

5.2 Real World Application Example

We verified our approach based on a real world application example, sense-through-foliage target detection from U.S. Air Force Research Laboratory. The target is a trihedral reflector (as shown in Fig. 4) in a forest. we plot two collections using UWB radars in Figs. 5a and 5b. Fig. 5a has no target on range, and Fig. 5b has target at samples around 13,900. We plot the echo differences between Figs. 5a and 5b in Fig. 5c. However, it is impossible to identify whether there is any target and where there is target based on Fig. 5c, which means *single radar doesn't work even in ideal case.* Since significant pulse-to-pulse variability exists in the echos, this motivate us to explore the spatial and time diversity using radar sensor networks approach. The echos, i.e., RF responses by the pulse of each cluster-member radar, are combined by the clusterhead using the RAKE structure in Fig. 1.

We ran simulations for an RSN with 30 radars, and plot the power of AC values in Figs. 6a and 6b for the two cases (with target and without target) respectively. Observe that in Fig. 6b, the power of AC values (around sample 13,900) where the target is located is non-fluctuating (monotonically increase then decrease). Although some other samples also have very high AC power values, it is very clear that they are quite fluctuating and the power of AC values behaves like random noise because generally the clutter has Gaussian distribution in the frequency domain.

6 Conclusions and Future Works

We have studied the constant frequency pulse waveform design and diversity in radar sensor networks. We showed that the waveforms can coexist if the carrier frequency spacing is a multiple of $1/T$ between two radar waveforms. We made analysis on interferences among waveforms in

RSN and proposed a RAKE structure for waveform diversity combining in RSN. As an application example, we applied the waveform design and diversity to CATR in RSN and proposed ML-CATR algorithms for non-fluctuating targets as well as fluctuating targets. Simulation results show that an RSN using our waveform diversity-based ML-ATR algorithms performs much better than a single radar system for non-fluctuating targets and fluctuating targets recognition. We also validated our RSN approach via a real-world sense-through-foliage application example.

In our future research, we will investigate the CATR when multiple targets coexist in RSN, and the number of targets are time-varying. In this paper, we used spatial diversity combining. For multi-target ATR, we will further investigate spatial-temporal-frequency combining for waveform diversity in RSN.

Acknowledgment

This work was supported by the U.S. Office of Naval Research (ONR) under Grant N00014-07-1-0395, N00014-07-1-1024, and the ONR Young Investigator Program Award under Grant N00014-03-1-0466.

References

[1] ONR BAA 06-016, "Command and Control and Combat Systems (C2 and CS)", http://www.onr.navy.mil/02/baa/expired.asp.

[2] ONR BAA 07-009, "Electronic Warfare Discovery and Invention (D&I)," http://www.onr.navy.mil/02/baa/.

[3] ONR BAA 07-017, "NET-SENTRIC Surveillance," http://www.onr.navy.mil/02/baa/.

[4] P. Baggenstoss, "Adaptive pulse length correction (APLECORR): a strategy for waveform optimization in ultrawideband active sonar," *IEEE Trans on Oceanic Engineering*, vol. 23, no. 1, pp. 1-11, 1998.

[5] C. E. Baum, et al, "The singularity expansion method and its application to target identification", *Proc. of the IEEE*, vol 79, no. 10, Oct 1991.

[6] M. R. Bell, "Information theory and radar waveform design", *IEEE Trans on Information Theory*, vol. 39, no. 5, pp. 1578-1597, Sept 1993.

[7] H. Deng, "Synthesis of binary sequences with good correlation and cross-correlation properties by simulated annealing," *IEEE Trans on Aerospace and Electronic Systems*, vol. 32, no. 1, Jan 1996.

[8] R. Fitzgerald, "Effects of range-Doppler coupling on chirp radar tracking accuracy," *IEEE Trans on Aerospace and Electronic Systems*, vol. 10, pp. 528-532, July 1974.

[9] T.-C. Hou and T-J. Tsai, " An access-based clustering protocol for multihop wireless ad hoc networks," *IEEE J. Selected Areas in Communications*, vol. 19, no. 7, pp. 1201-1210, July 2001.

[10] A. Iwata, C. C. Chiang, G. Pei, M. Gerla, and T. W. Chen, " Scalable routing strategies for ad hoc networks," *IEEE J. Selected Areas in Communications*, vol. 17, pp. 1369-1379, 1999.

[11] R. A. Johnson and E. L. Titlebaum, "Range Doppler Uncoupling in the Doppler Tolerant Bat Signal", *Proc. of IEEE Ultrasonics Symposium*, New York, pp. 64-67, 1972.

[12] D. Kershaw and R. Evans, "Optimal waveform selection for tracking system", *IEEE Trans on Information Theory*, vol. 40, no. 5, pp. 1536-1550, 1994.

[13] C. R. Lin and M. Gerla, "Adaptive clustering in mobile wireless networks," *IEEE J. Selected Areas in Communications*, vol. 16, pp. 1265-1275, 1997.

[14] Q. Liang, X. Cheng, "KUPS: Knowledge-based Ubiquitous and Persistent Sensor Networks for Threat Assessment", accepted by *IEEE Trans on Aerospace and Electronic Systems.*

[15] R. Niu, P. Willett, and Y. Bar-Shalom, "Tracking consideration in selection of radar waveform for range and range-rate measurements", *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 38, No. 2, 2002.

[16] A. Papandreou, G. F. Boudreaux-Bartels, and S. M. Kay, "Detection and estimation of generalized chirps using time-frequency representation", *Twenty-Eighth Asilomar Conference on Signals, Systems and Computers*, vol. 1, pp. 50-54, Oct. 1994.

[17] C. E. Perkins, "Chapter 4, Cluster-Based Networks," *Ad Hoc Networking*, Edited by C. E. Perkins, pp. 75-138, Addison-Wesley, 2001.

[18] M. A. Richards, *Fundamentals of Radar Signal Processing*, McGraw-Hill Companies, New York, 2005.

[19] J. Roman, M. Rangaswamy, D. Davis, Q. Zhang, B. Himed, and J. Michels, "Parametric adaptive matched filter for airborne radar applications," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 36, no. 2, pp. 677-692, 2000.

[20] T. K. Sarkar and N. Sangruji, "An adaptive nulling system for a narrow-band signal with a look-direction constraint utilizing the conjugate gradient method," *IEEE Trans. Antennas Propagat.*, vol. 37, no. 7, pp. 940-944, July 1989.

[21] M. I. Skolnik, *Introduction to Radar Systems*, 3rd ed, New York, McGraw Hill, 2001.

[22] P. Swerling, "Probability of detection for fluctuating targets", *IRE Trans on Information Theory*, vol. 6, pp. 269-308, April 1960.

[23] Y. Sun, P. Willett, and R. Lynch, "Waveform fusion in sonar signal processing", *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 40, No. 2, 2004

[24] S. Sowelam and A. Tewfik, "Waveform selection in radar target classification," *IEEE Trans on Information Theory*, vol. 46, no. 3, pp. 1014-1029, 2000.

List of Tables

Table 1: RCS values at microwave frequency for 5 targets.

Index n	Target	RCS
1	Bird	0.01
2	Conventional unmanned winged missile	0.5
3	Small single-engine aircraft	1
4	Small fighter aircraft or 4 passenger jet	2
5	Large fighter aircraft	6

List of Figures

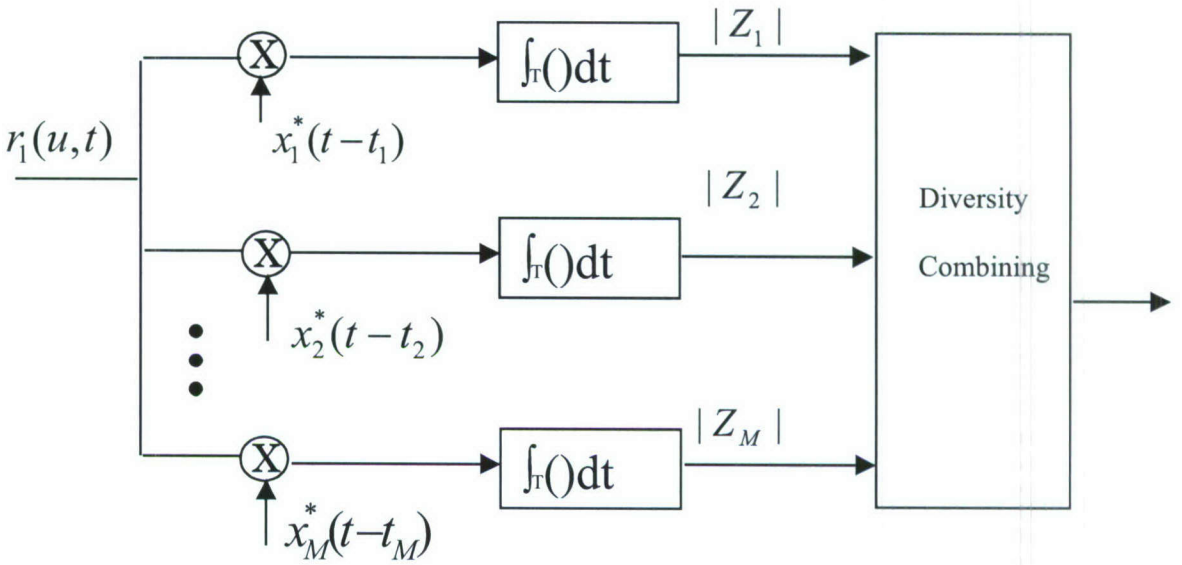
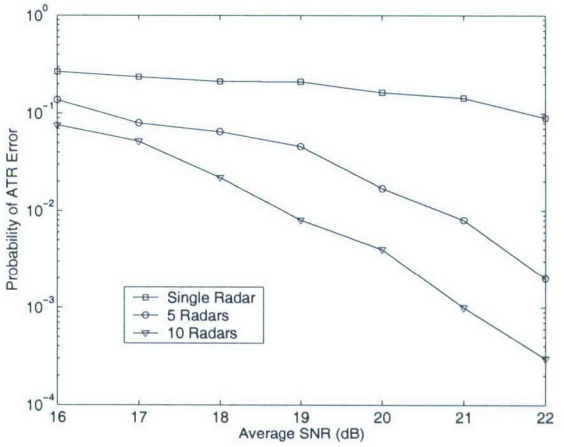
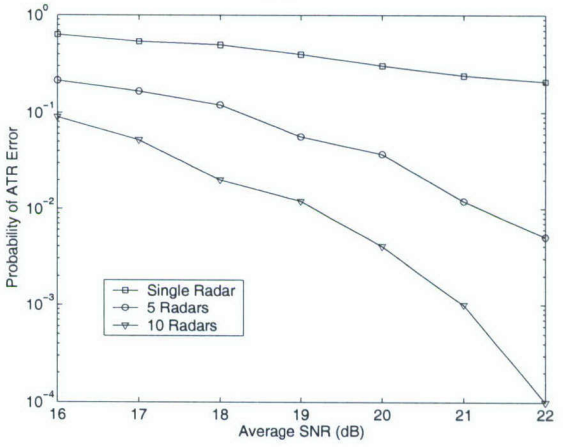


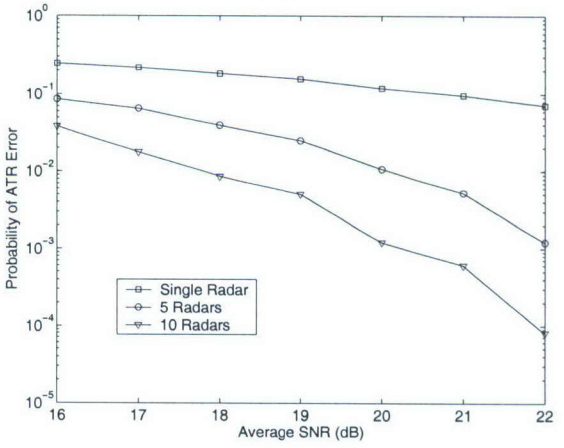
Figure 1: Waveform diversity combining by clusterhead in RSN.



(a)

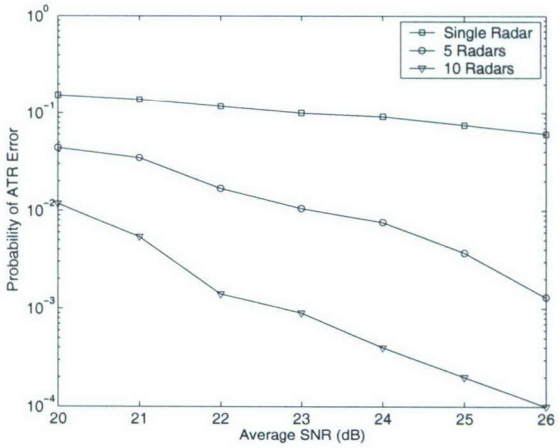


(b)

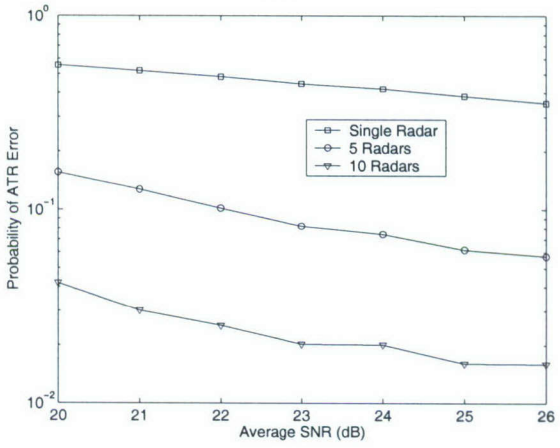


(c)

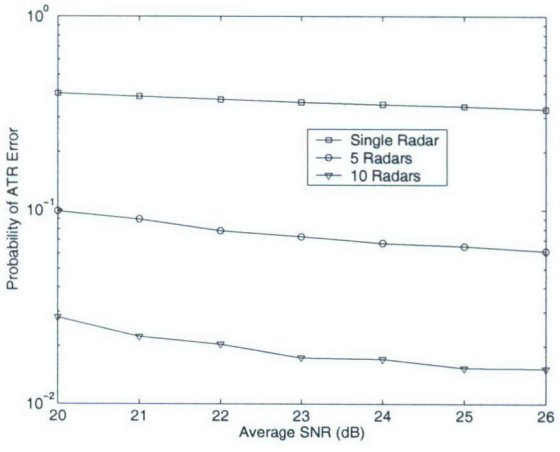
Figure 2: Probability of ATR error for *non-fluctuating* targets at different average SNR (dB) values. (a) bird, (b) missile, (c) the average probability of ATR error for 5 targets.



(a)



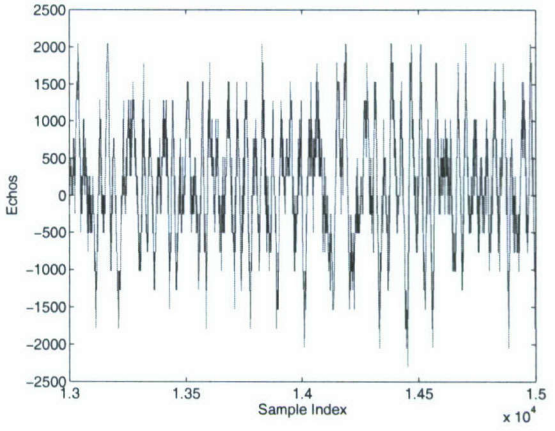
(b)



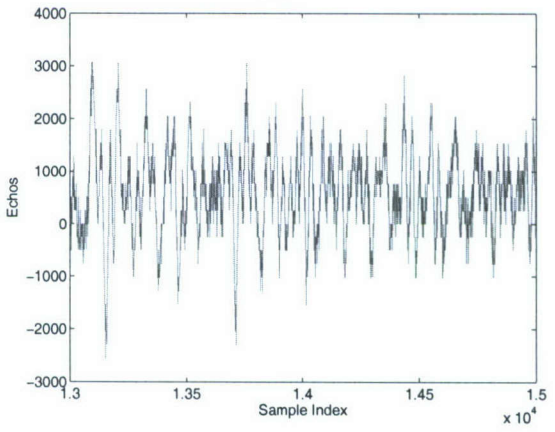
(c)

Figure 3: Probability of ATR error for *fluctuating* targets at different average SNR (dB) values. (a) bird, (b) missile, (c) the average probability of ATR error for 5 targets.

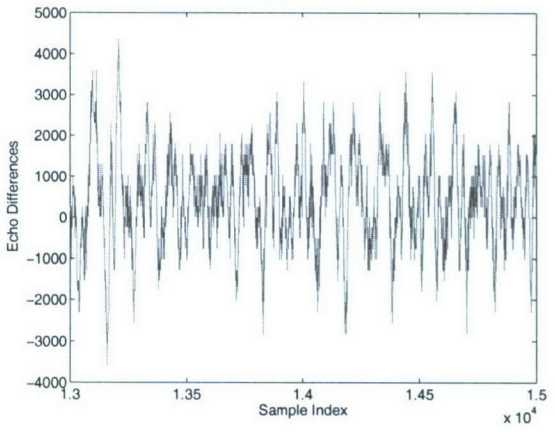
Figure 4: The target (a trihedral reflector) is shown on the stand at 300 feet from the measurement lift.



(a)

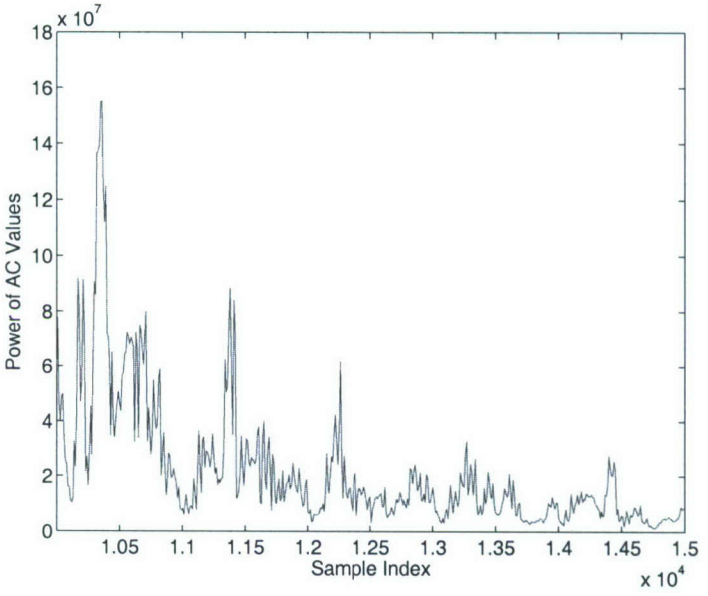


(b)

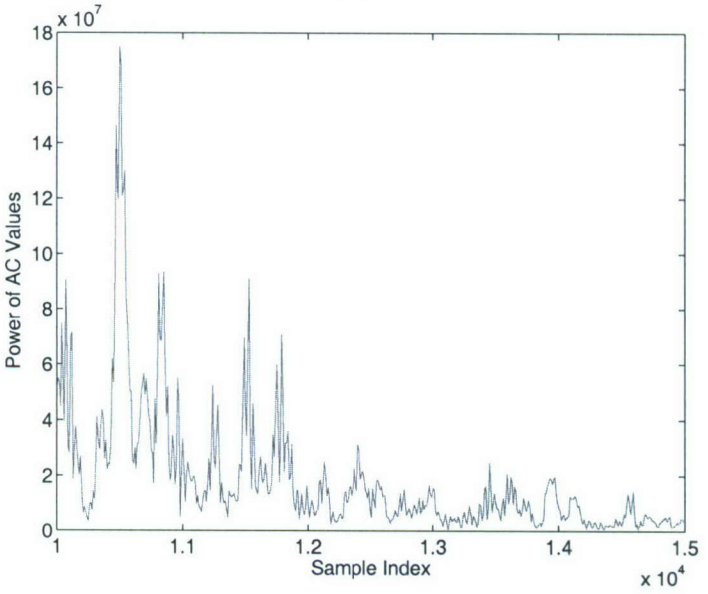


(c)

Figure 5: Measurement of UWB radar. (a) Expanded view of traces (no target) from sample 13,001 to 15,000. (b) Expanded view of traces (with target) from sample 13,001 to 15,000. (c) The differences between (a) and (b).



(a)



(b)

Figure 6: Power of AC values based on radar sensor networks approach. (a) No target (b) With target in the field (in samples around 13,900).

Foliage Clutter Modeling Using Narrowband and UWB Radars

Jing Liang, Qilian Liang
Department of Electrical Engineering
University of Texas at Arlington
Arlington, TX 76019-0016, USA
Email: jliang@wcn.uta.edu, liang@uta.edu

Abstract— **In this paper, we study foliage clutter modeling on a basis of both pragmatically narrowband and ultra-wide band (UWB) radars. We propose that the foliage clutter follows log-logistic model using maximum likelihood (ML) parameter estimation as well as the root mean square error (RMSE) on PDF curves between original clutter and statistical model data. In addition to investigating the log-logistic model, we also compare it with other popular clutter models, namely log-normal and Weibull. We show that the log-logistic model not only achieves the smallest standard deviation (STD) error on estimated model parameters, but also has the best goodness-of-fit and smallest RMSE.**

I. Introduction and Motivation

Clutter is a term used to define all unwanted echoes from natural environment [1]. The nature of clutter may necessarily vary on a basis of different applications and radar parameters. Most previous studies have investigated land clutter and sea clutter. For example, log-normal and Weibull distributions have been proven to be better suited for the clutter than Rayleigh and Rician models in high resolution radar systems. As far as clutter modeling in forest is concerned, it is still of great interest and is likely to take some time to reach any agreement. A team of researchers from MIT [2] and U. S. Army Research Laboratory (ARL) [3] [4] have measured ultra-wideband (UWB) backscatter signals in foliage for different polarizations and frequency ranges. The measurements show that the foliage clutter is impulsively corrupted with multi-path fading, which leads to inaccuracy of the K-distributions description [5]. The Air Force Office of Scientific Research (AFOSR) has conducted field measurement experiment concerning foliage penetration radar since 2004 and noted that metallic targets may be more easily identified with wideband than with narrowband signals [6].

In this investigation, we will apply both narrowband and ultra-wide band (UWB) radar to model the foliage clutter, as we believe that foliage clutter is composed of intervening materials that are electromagnetically dispersive, which contributes to the strong frequency dependence of foliage, and thus a narrowband-wideband study would assist with the better understanding of statistic property of the clutter.

Narrowband signals have been tried at 200, 400 and 600 megahertz respectively, while UWB radar emissions are at a relatively low frequency-typically between 100 MHz and 3 GHz. Each frequency component in a radar signal will sense the foliage clutter in a slightly different manner, therefore provide differences in multipath.

In our present work, we investigate the use of the log-logistic distribution (LLD) to model foliage clutter and illustrate the goodness-of-fit to real UWB clutter data. Additionally, we compare the goodness-of-fit with existing popular models namely log-normal, Weibull by means of maximum likelihood estimation (MLE) and the root mean square error (RMSE). The result shows that log-logistic model provides the best fit to the foliage clutter for both narrowband and UWB signals.

The rest of this paper is organized as follows. In Section II we discuss the properties and applicability of log-logistic as a statistical model for foliage clutter. The measurement and collection of clutter data we used in this paper are summarized in Section III. Section IV discribes the estimation of model parameters and the goodness-of-fit. Finally, section V concludes this paper and describes some future work.

II. Log-logistic Model

Log-logistic has been applied recently in hydrological analysis. In spite of its intensive application in precipitation and stream-flow data, the LLD [7] statistical model, to the best of our knowledge, has never been applied to radar foliage clutter. The motivation for considering log-logistic model is based on its higher kurtosis and longer tails, as well as its shape similarity to log-normal and Weibull distributions. Thus it is intended to be employed to estimate how well the model matches our collected foliage clutter statistics.

The PDF for LLD distribution is given by

$$f(x) = \frac{e^{\frac{\ln x-\mu}{\sigma}}}{\sigma x(1+e^{\frac{\ln x-\mu}{\sigma}})^2}, \quad x>0,\ \sigma>0 \tag{1}$$

where μ is scale parameter and σ is shape parameter.

The mean of the the LLD is

$$E\{x\} = e^{\mu}\Gamma(1+\sigma)\Gamma(1-\sigma) \tag{2}$$

whereas the median is

$$M\{x\} = e^{\mu} \tag{3}$$

The variance is given by

$$Var\{x\} = e^{2\mu}\{\Gamma(1+2\sigma)\Gamma(1-2\sigma) - [\Gamma(1+\sigma)\Gamma(1-\sigma)]^2\} \tag{4}$$

while the moment of order k is

$$E\{x^k\} = \sigma e^{\mu} B(k\sigma, 1 - k\sigma), \; k < \frac{1}{\sigma} \tag{5}$$

where

$$B(m,n) = \int_0^1 x^{m-1}(1-x)^{n-1} dx \tag{6}$$

The PDFs for selected μ's and σ's for the LLD distribution are illustrated in Fig. 1. We will apply LLD along with lognormal and Weibull models to analyze the clutter data.

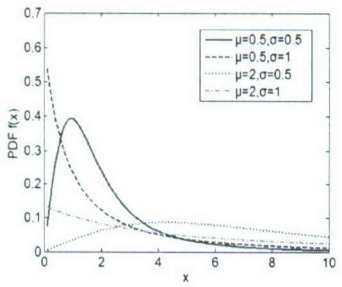


Fig. 1. Log-logistic distribution PDF for $\mu = 0.5$ and $\sigma = 0.5$, $\mu = 0.5$ and $\sigma = 1$, $\mu = 2$ and $\sigma = 0.5$, $\mu = 2$ and $\sigma = 1$

III. Experiment Setup and Data Collection

The foliage penetration measurement effort began in August 2005 and continued through December 2005. Working in August through the fall of 2005, the foliage measured included late summer foliage and fall and early winter foliage. Late summer foliage, because of the limited rainfall, involved foliage with decreased water content. Late fall and winter measurements involved largely defoliated but dense forest.

Fig. 2. This figure shows the lift with the experiment. The antennas are at the far end of the lift from the viewer under the roof that was built to shield the equipment from the elements. This picture was taken in September with the foliage largely still present. The cables coming from the lift are a ground cable to an earth ground and one of 4 tethers used in windy conditions.

The radar experiment was constructed on a seven-ton man lift, which had a total lifting capacity of 450 kg. The limit of the lifting capacity was reached during the experiment as essentially the entire measuring apparatus was placed on the

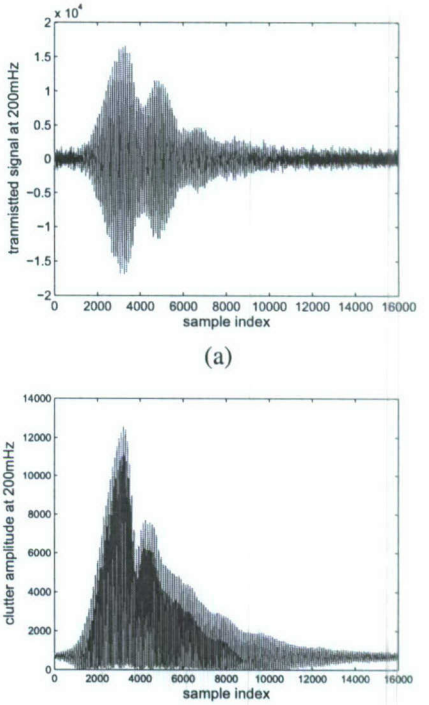


Fig. 3. Signal at 200 megahertz (a) transmitted pulse (b) received pulse

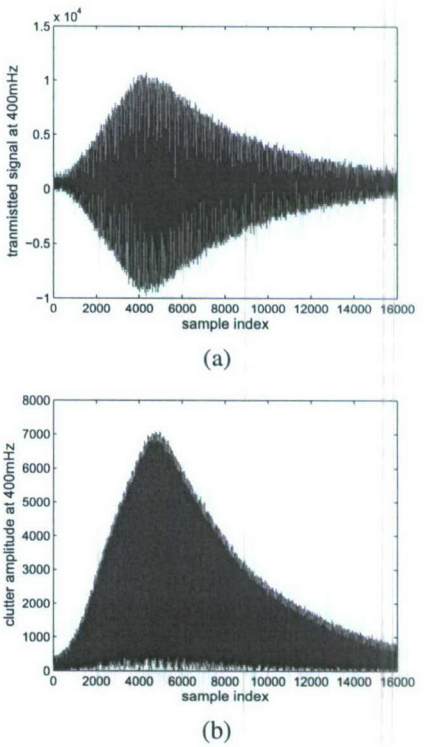
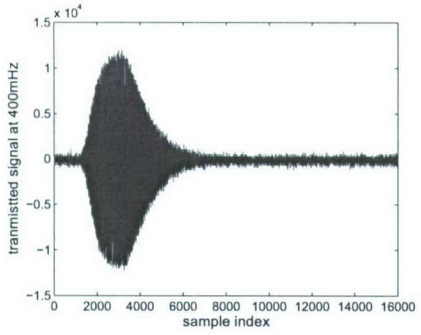
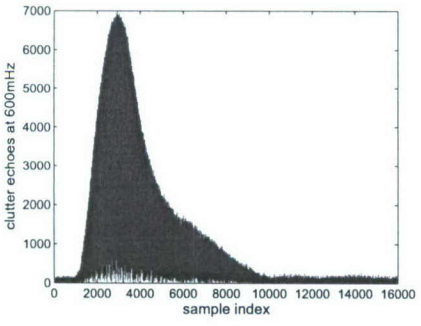


Fig. 4. Signal at 400 megahertz (a) transmitted pulse (b) received pulse

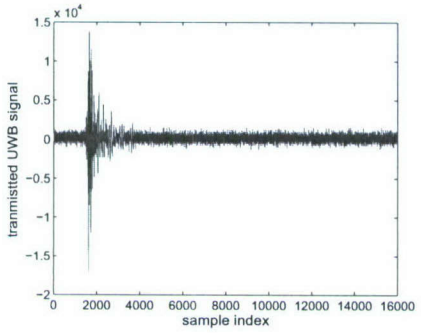


(a)

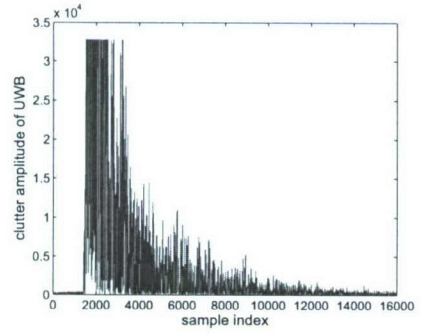


(b)

Fig. 5. Signal at 600 megahertz (a) transmitted pulse (b) received pulse



(a)



(b)

Fig. 6. Signal at 600 megahertz (a) transmitted pulse (b) received pulse

lift (as shown in Fig. 2). The principle pieces of equipment secured on the lift are: dual antenna mounting stand; two antennas; 200MHz, 1Kw Amplifier, power supply, pre-amp; 400MHz, 1Kw Amplifier, power supply, pre-amp; 600MHz, 1Kw Amplifier, power supply, pre-amp; Tektronix model 7704 B oscilloscope; rack system; IBM laptop; HP signal Generator; Custom RF switch and power supply and weather shield (small hut). Particularly for UWB signal, a Barth pulse source (Barth Electronics, Inc. model 732 GL) was used. The pulse generator uses a coaxial reed switch to discharge a charge line for a very fast rise time pulse outputs. The model 732 pulse generator provides pulses of less than 50 picoseconds (ps) rise time, with amplitude from 150 V to greater than 2 KV into any load impedance through a 50 ohm coaxial line. The generator is capable of producing pulses with a minimum width of 750 ps and a maximum of 1 microsecond. This output pulse width is determined by charge line length for rectangular pulses, or by capacitors for 1/e decay pulses.

The system was pointing at the specified 250 feet one way distance. For the data we used in this paper, each sample is spaced at 50 picoseconds interval, and 16,000 samples were collected for each collection for a total time duration of 0.8 microseconds. 35 pulses reflected clutter signal were obtained for each collection at the same frequency and the same site, but different time.

The transmitted pulses and received signal are shown in Fig. 3$\sim$ 6. We averaged the 35 pulses in order to remove the random noise from the clutter. It is obvious to see the presence of multiple scattering from received signals, UWB in particular.

IV. Statistical Analysis of the Foliage Clutter Data

A. Maximum Likelihood Estimation

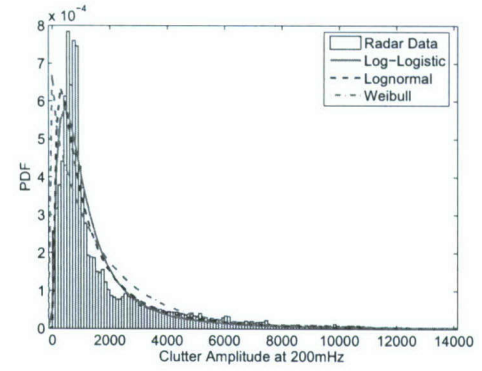
Using the collected clutter data mentioned above, we apply Maximum Likelihood Estimation (MLE) approach to estimate the parameters of the log-logistic, log-normal and Weibull models. MLE is often used when the sample data are known and parameters of the underlying probability distribution are to be estimated [8] [9]. It is generalized as follows:

Let $y_1, y_2, \cdots, y_N$ be N independent samples drawn from a random variable $\mathbf{Y}$ with m parameters $\theta_1, \theta_2, \cdots, \theta_m$, where $\theta_i \in \theta$, then the likelihood function expressed as a function of θ conditional on $\mathbf{Y}$ is

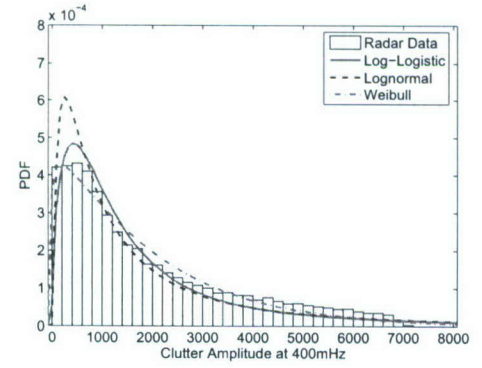
$$L_N(\mathbf{Y}|\theta) = \prod_{k=1}^{N} f_{Y|\theta}(y_k|\theta_1, \theta_2, \cdots, \theta_m) \tag{7}$$

The maximum likelihood estimate of $\theta_1, \theta_2, \cdots, \theta_m$ is the set of values $\hat{\theta}_1, \hat{\theta}_2, \cdots, \hat{\theta}_m$ that maximize the likelihood function $L_N(\mathbf{Y}|\theta)$.

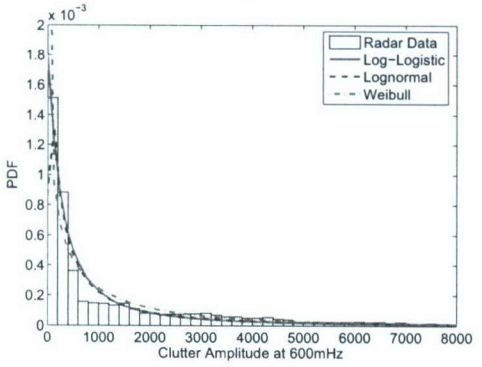
As the logarithmic function is monotonically increasing, maximizing $L_N(\mathbf{Y}|\theta)$ is equivalent to maximizing $\ln(L_N(\mathbf{Y}|\theta))$. Hence, it can be shown that a necessary but not sufficient condition to obtain the ML estimate $\hat{\theta}$ is to solve the



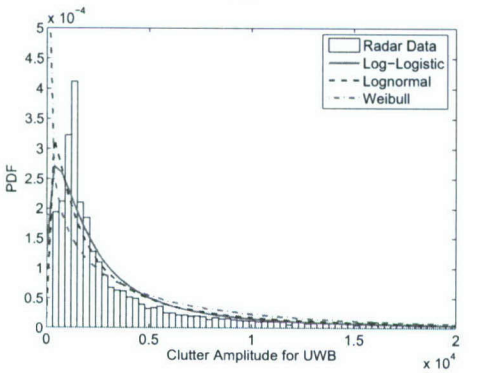
(a)



(b)



(c)



(d)

Fig. 7. Model Comparison (a) 200 megahertz (b) 400 megahertz (c) 600 megahertz (d) UWB

TABLE I
ESTIMATED PARAMETERS AND STD ERROR

PDF	Log-Logistic	Log-normal	Weibull
200MHz	$\hat{\mu} = 6.95521$ $\hat{\sigma} = 0.619019$ $\varepsilon_\mu = 0.00848$ $\varepsilon_\sigma = 0.00409$	$\hat{\mu} = 6.95754$ $\hat{\sigma} = 1.12545$ $\varepsilon_\mu = 0.00890$ $\varepsilon_\sigma = 0.00629$	$\hat{a} = 1816.39$ $\hat{b} = 0.95900$ $\varepsilon_a = 15.8698$ $\varepsilon_b = 0.00560$
400MHz	$\hat{\mu} = 7.138$ $\hat{\sigma} = 0.67077$ $\varepsilon_\mu = 0.00922$ $\varepsilon_\sigma = 0.00441$	$\hat{\mu} = 7.03467$ $\hat{\sigma} = 1.22477$ $\varepsilon_\mu = 0.00968$ $\varepsilon_\sigma = 0.00685$	$\hat{a} = 1970.62$ $\hat{b} = 1.06597$ $\varepsilon_a = 15.38150$ $\varepsilon_b = 0.00671$
600MHz	$\hat{\mu} = 6.26835$ $\hat{\sigma} = 0.98842$ $\varepsilon_\mu = 0.01334$ $\varepsilon_\sigma = 0.00637$	$\hat{\mu} = 6.24074$ $\hat{\sigma} = 1.68841$ $\varepsilon_\mu = 0.01383$ $\varepsilon_\sigma = 0.00944$	$\hat{a} = 1164.35$ $\hat{b} = 0.67637$ $\varepsilon_a = 14.3965$ $\varepsilon_b = 0.00414$
UWB	$\hat{\mu} = 7.76868$ $\hat{\sigma} = 0.78651$ $\varepsilon_\mu = 0.01078$ $\varepsilon_\sigma = 0.00522$	$\hat{\mu} = 7.79566$ $\hat{\sigma} = 1.41771$ $\varepsilon_\mu = 0.01121$ $\varepsilon_\sigma = 0.00793$	$\hat{a} = 4901.07$ $\hat{b} = 0.74322$ $\varepsilon_a = 55.3011$ $\varepsilon_b = 0.00434$

likelihood equation

$$\frac{\partial}{\partial \theta} \ln(L_N(\mathbf{Y}|\theta)) = 0 \tag{8}$$

Using the collected clutter radar mentioned above, we apply MLE to obtain $\hat{\mu}$ and $\hat{\sigma}$ for log-logistic, $\hat{\mu}$ and $\hat{\sigma}$ for the log-normal, $\hat{a}$ and $\hat{b}$ for the Weibull. The results are shown in Table I. We also explore the standard deviation (STD) error of each parameter. These descriptions are also shown in table I in the form of ε_x, where x denotes different parameter for each model.

From Table I, we can see that no matter what frequency is used, in general STD error for log-logistic parameters are smaller than those of log-normal and Weibull.

B. Goodness-of-fit in curve and RMSE

We may also observe the extend to which the PDF curve of the statistic model matches that of clutter data by intuitive visual inspection as well as calculating the root mean square error (RMSE).

The goodness-of-fit in curve are illustrated in Fig. 7. The histograms of original radar clutter data are skewed to the right. In general, LLD tends to have lower kurtosis, steeper slope than those of lognormal and Weibull but lower PDF in tail than that of lognormal. At 200MHz, LLD and lognormal are superior to Weibull. At 400MHz, the performance of lognormal is the worst. At 600MHz, Weibull become the worst again and as for UWB signal, Weibull can not fit well.

To further evaluate the goodness-of-fit, the RMSE is applied. Let i (i=1, 2, $\cdots$, n) be the sample index of clutter amplitude, c_i is the corresponding PDF value whereas $\hat{c}_i$ is the PDF value of the statistical model with estimated parameters by means of MLE as mentioned above. The RMSE is obtained through the following equation:

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(c_i - \hat{c}_i)^2} \tag{9}$$

Here we apply n=101 for each model. The result is given in Table II.

TABLE II
ROOT MEAN SQUARE ERROR (RMSE)

PDF	Log-Logistic	Log-normal	Weibull
200MHz	2.157×10^{-5}	2.3248×10^{-5}	3.5849×10^{-5}
400MHz	2.2423×10^{-5}	2.8956×10^{-5}	2.6764×10^{-5}
600MHz	3.3864×10^{-5}	3.4278×10^{-5}	3.7259×10^{-5}
UWB	2.739×10^{-5}	3.1866×10^{-5}	4.4045×10^{-5}

This shows that despite of visual inspection, the log-logistic model is more accurate than the other two models.

V. CONCLUSION

On a basis of foliage clutter data measured by the both narrowband (200MHz, 400MHz and 600MHz) and UWB radars, we show that it is more accurate to describe foliage clutter using log-logistic statistic model rather than log-normal, Weibull. Future research will investigate the characteristics of target to better achieve the target detection, tracking and imaging.

ACKNOWLEDGEMENT

This work was supported in part by Office of Naval Research (ONR) under Grant N00014-07-1-0395, N00014-07-1-1024. The authors would like to thank Dr. Sherwood Samn in AFRL for providing the UWB radar data.

REFERENCES

[1] Skolnik, M. I.,*Introduction to Radar Systems*, 3rd ed, New York, McGraw Hill, 2001.

[2] Fleischman, J. G., Ayasli, S., Adams, E. M., and Gosselin, D. R., Foliage penetration experiment: Part I: Foliage attenuation and backscatter analysis of SAR imagery, *IEEE Trans. on Aerosp. Electron. Syst.*, 32, 1, part 1 of 3 (1996), 134-144.

[3] Mccorkle, J. W., "Early results from the ARL UWB Foliage attenuation (FOPEN) SAR", Presented at the SPIE International Symposium on Optical Engineering and Photonics in Aerospace and Remote Sensing, Conference 1942, Underground and Obscured Object Detection, Apr. 1993.

[4] Sheen, D. R., Malinas, N. P., Kletzli, D. W., Lewis, T. B., and Roman, J. F., "Foliage transmission meassurement using a ground-based ultra-wideband (UWB) (300-1300MHz) SAR system", *IEEE Trans. on Geoscience and Remote Sensing*, 32, 1(1994).

[5] Watts, S., "Radar detection prediction in K-distribution sea clutter and thermal noise", *IEEE Trans. on Aerosp. Electron. Syst.*, AES-23, 1, 1987, pp. 40-45.

[6] Dill, C., "Foliage Penetration (Phase II) Field Test: Narrowband versus Wideband Foliage Penetration," Final Report of Contract Number F41624-03-D-7001/04, July 2005 to Feb 2006.

[7] Gupta, R. C., Akman, O. and Lvin, S., "A Study of Log-Logistic Model in Survival Analysis", *Biometrical Journal*, 41, pp. 431-443, 1999

[8] Devore, Probability and Statistics for Engineering and the Sciences. Monterey, CA: BrooksiCole, 1982.

[9] Barkat, M., Signal deteciton and estimation, 2nd, London: Artech house, 2005.

A Propagation Environment Modeling in Foliage

Jing Liang and Qilian Liang, *Senior Member, IEEE*
Department of Electrical Engineering
University of Texas at Arlington
Arlington, TX 76019-0016, USA
E-mail: jliang@wcn.uta.edu, liang@uta.edu

Sherwood W. Samn
Air Force Research Laboratory/HEX
Brooks City Base
San Antonio, TX 78235, USA
Sherwood.samn@brooks.af.mil

Abstract

In this paper, we propose that foliage clutter follows log-logistic model using maximum likelihood (ML) parameter estimation as well as the root mean square error (RMSE) on PDF curves between original clutter and statistical model data. We not only investigate log-logistic model, but compare it with other popular clutter models, namely log-normal, Weibull and Nakagami. It shows that the log-logistic model not only achieves the smallest standard divination (STD) error on estimated model parameters, but also the best goodness-of-fit and smallest RMSE for both poor and good clutter signals.

Index Terms : foliage clutter, log-logistic, log-normal, Weibull, Nakagami, goodness-of-fit

1 Introduction and Motivation

Detection and identification of military equipment in a strong clutter background, such as foliage, soil cover or building has been a long-standing subject of intensive study. It is believed that solving the target detection through foliage environment will significantly benefit sense-through-wall and many other subsurface sensing problems. However, to this date, the detection of foliage-covered military targets with the required probability of detection and false alarm still remains a challenging issue. Recent investigations on environment behavior of tree canopies

have shown that both signal backscattering and attenuation are significantly influenced by tree architecture [2]. Therefore use the return signal from foliage to establish the clutter model that accounts for environment effects is crucial for the sense-through-foliage radar detection.

Clutter is a term used to define all unwanted echoes from natural environment [3]. The nature of clutter may necessarily vary on a basis of different applications and radar parameters. Most previous studies have investigated land clutter or sea clutter, and some conclusions have been reached. For example, log-normal, Weibull, and K-distributions have been proven to be better suited for the clutter description other than Rayleigh and Rician models in high resolution radar systems. Fred [4] did statistical comparisons and found that sea clutter at low grazing angles and high range resolution is spiky based on the data measured from various sites in Kauai and Hawaii. David generalized radar clutter models using noncentral chi-square density by allowing the noncentrality parameter to fluctuate according to the gamma distribution [5]. Furthermore, Henry *et al.* used a Neural-Network-based approach to predict sea clutter model [6] [7].

However, as far as clutter modeling in forest is concerned, it is still of great interest and will be likely to take some time to reach any agreement. A team of researchers from MIT [8] and U. S. Army Research Laboratory (ARL) [9] [10] have measured ultra-wideband (UWB) backscatter signals in foliage for different polarizations and frequency ranges. The measurements show that the foliage clutter is impulsively corrupted with multipath fading, which leads to inaccuracy of the K-distributions description [11]. The Air Force Office of Scientific Research (AFOSR) has conducted field measurement experiment concerning foliage penetration radar since 2004 and noted that metallic targets may be more easily identified with wideband than with narrowband signals [12].

In this investigation, we will apply ultra-wide band (UWB) radar to model the foliage clutter. UWB radar emissions are at a relatively low frequency-typically between 100 MHz and 3 GHz. Additionally, the fractional bandwidth of the signal is very large (greater than 0.2). Such a radar sensor has exceptional range resolution that also has an ability to penetrate

many common materials (e.g., walls). Law enforcement personnel have used UWB ground penetrating radars (GPRs) for at least a decade. Like the GPR, sense-through-foliage radar takes advantage of UWB's very fine resolution (time gating) as well as low frequency of operation.

In our present work, we investigate the use of the log-logistic distribution to model foliage clutter and illustrate the goodness-of-fit to real UWB clutter data conducted by AFOSR. Additionally, we compare the goodness-of-fit with existing popular models namely log-normal, Weibull, and Nakagami by means of maximum likelihood estimation (MLE) and the root mean square error (RMSE). The result shows that log-logistic model provides the best fit to the foliage clutter. Our contribution is not only the new proposal on the foliage clutter model with estimated concrete parameters, but also provide the criteria and approaches based on which the statistical analysis is deduced. Further, the theoretical study about the probability of detection as well as the probability of false alarm is discussed.

The rest of this paper is organized as follows. Section 2 provides a statistical model review on log-logistic, log-normal, Weibull and Nakagami distributions and discuss their properties and applicability as models for foliage clutter. Section 3 summarizes the measurement and the 2 sets of clutter data that we used in this paper. Section 4 discusses estimation on parameters and the goodness-of-fit for log-logistic, log-normal, Weibull and Nakagami models respectively. Section 5 analyzes the performance of radar detection at presence of foliage clutter. Finally, section 6 concludes this paper and describes some future research topics.

2 Clutter Models

Many radar clutter models have been proposed in terms of distinct statistical distributions, most of which describe the characteristics of clutter amplitude or power. Before detailed analysis, first we would like to discuss the properties and applicability of log-logistic, log-normal, Weibull, and Nakagami statistic distributions, which are designated as "curve fit" models in section 4, since they are more likely to provide good fit to our collections of pragmatic

clutter data in general. Detailed explanations would be given in following subsections.

2.1 Log-logistic Model

Recently Log-logistic model has been applied in hydrological analysis. This distribution is a special case of Burr's type-XII distribution [15] as well as a special case of the kappa distribution proposed by Mielke and Jonson [16]. Lee *et al.* employed the LLD for frequency analysis of multiyear drought durations [17], whereas Shoukri *et al.* employed LLD to analyse extensive Canadian precipitation data [18], and Narda & Malik used LLD to develop a model of root growth and water uptake in wheat [19]. In spite of its intensive application in precipitation and stream-flow data, the log-logistic distribution (LLD) [13] statistical model, to the best of our knowledge, has never been applied to radar foliage clutter. The motivation for considering log-logistic model is based on its higher kurtosis and longer tails, as well as its PDF curve similarity to log-normal and Weibull distributions. It is intended to be employed to estimate how well the model matches our collected foliage clutter statistics.

Here we apply the two-parameter distribution with parameters μ and σ. The PDF for this distribution is given by

$$f(x) = \frac{e^{\frac{\ln x-\mu}{\sigma}}}{\sigma x(1+e^{\frac{\ln x-\mu}{\sigma}})^2}, \quad x>0,\ \sigma>0 \tag{1}$$

where μ is scale parameter and σ is shape parameter. The mean of the the LLD is

$$E\{x\} = e^{\mu}\Gamma(1+\sigma)\Gamma(1-\sigma) \tag{2}$$

The variance is given by

$$Var\{x\} = e^{2\mu}\{\Gamma(1+2\sigma)\Gamma(1-2\sigma) - [\Gamma(1+\sigma)\Gamma(1-\sigma)]^2\} \tag{3}$$

while the moment of order k is

$$E\{x^k\} = \sigma e^{\mu} B(k\sigma, 1-k\sigma),\ k<\frac{1}{\sigma} \tag{4}$$

where

$$B(m,n) = \int_0^1 x^{m-1}(1-x)^{n-1}dx \tag{5}$$

PDFs for LLD for selected μ's and σ's are illustrated in Fig. 1.

2.2 Log-normal Model

Most previous experimental data have resulted in clutter being modeled using a log-normal distribution, which is most frequently used when the radar sees land clutter [20] or sea clutter [21] at low grazing angles (≤ 5 degrees) since log-normal has a long tail. However, it has been reported that the log-normal model tends to overestimate the dynamic range of the real clutter distribution [22]. Furthermore, most previous research applies log-normal model to land and sea clutter, but how accurately it models foliage clutter requires detailed analysis.

The log-normal distribution [23] is also a two-parameter distribution with parameters μ and σ. The PDF for this distribution is given by

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}}, \quad x > 0, \sigma > 0 \tag{6}$$

where μ is the scale parameter and σ is the shape parameter. The mean, variance and the moment of order k are given respectively by

$$E\{x\} = e^{\mu + \frac{\sigma^2}{2}} \tag{7}$$

$$Var\{x\} = (e^{\sigma^2} - 1)e^{2\mu + \sigma^2} \tag{8}$$

$$E\{x^k\} = e^{k\mu + \frac{k^2\sigma^2}{2}} \tag{9}$$

PDFs for selected μ's and σ's for log-normal distribution are shown in Fig. 2.

2.3 Weibull Model

The Weibull distribution, which is named after Waloddi Weibull, can be made to fit clutter measurements that lie between the Rayleigh and log-normal distribution [24]. It has been applied to land clutter [25] [26], sea clutter [27] [28] and weather clutter [29]. However, in

very spiky sea and foliage clutter, the description of the clutter statistics provided by Weibull distributions may not always be sufficiently accurate [30].

The Weibull distribution is also a two-parameter distribution with parameters a and b. The PDF for this distribution is given by

$$f(x) = ba^{-b}x^{b-1}e^{-(x/a)^b}, \quad x > 0, a > 0, \ b > 0 \tag{10}$$

where b is the shape parameter and a is the scale parameter. The mean, variance and the moment of order k are given respectively by

$$E\{x\} = a\Gamma(1 + \frac{1}{b}) \tag{11}$$

$$Var\{x\} = a^2\{\Gamma(1 + \frac{2}{b}) - [\Gamma(1 + \frac{1}{b})]^2\} \tag{12}$$

$$E\{x^k\} = a^k\Gamma(1 + \frac{k}{b}) \tag{13}$$

PDFs for selected a's and b's for Weibull distribution are shown in Fig. 3.

2.4 Nakagami Model

In the foliage penetration setting, the target returns suffer from multipath effects corrupted with fading. As Nakagami distribution is used to model scattered fading signals that reach a receiver by multiple paths, it is natural to investigate how well it fits the foliage clutter statistics.

The PDF for Nakagami distribution is given by

$$f(x) = 2(\frac{\mu}{\omega})^{\mu}\frac{1}{\Gamma(\mu)}x^{(2\mu-1)}e^{-\frac{\mu}{\omega}x^2}, \quad x > 0, \omega > 0 \tag{14}$$

where μ is the shape parameter and ω is the scale parameter. The mean, variance and the moment of order k of Nakagami distribution are given respectively by

$$E\{x\} = \frac{\Gamma(\mu + \frac{1}{2})}{\Gamma(\mu)}(\frac{\omega}{\mu})^{\frac{1}{2}} \tag{15}$$

$$Var\{x\} = \omega[1 - \frac{1}{\mu}(\frac{\Gamma(\mu + \frac{1}{2})}{\Gamma(\mu)})^2] \tag{16}$$

$$E\{x^k\} = \frac{\Gamma(\mu + \frac{k}{2})}{\Gamma(\mu)}(\frac{\omega}{\mu})^{\frac{k}{2}} \tag{17}$$

The PDFs for selected μ's and ω's for the Nakagami distribution are illustrated in Fig. 4.

3 Experiment Setup and Data Collection

Our work is based on the sense-through-foliage data from Air Force Research Lab [12]. The foliage penetration measurement effort began in August 2005 and continued through December 2005. Working in August through the fall of 2005, the foliage measured included late summer foliage and fall and early winter foliage. Late summer foliage, because of the limited rainfall, involved foliage with decreased water content. Late fall and winter measurements involved largely defoliated but dense forest.

A bistatic UWB radar (individual transmit and receive antennas) system was used and the experiment was constructed on a seven-ton man lift, which had a total lifting capacity of 450 kg. The limit of the lifting capacity was reached during the experiment as essentially the entire measuring apparatus was placed on the lift (as shown in Fig. 5). The principle pieces of equipment secured on the lift are: Barth puls generator, Tektronix model 7704 B oscilloscope, dual antenna mounting stand, two antennas, rack system, IBM laptop, HP signal Generator, Custom RF switch and power supply and Weather shield (small hut). Throughout this work, a Barth pulse source (Barth Electronics, Inc. model 732 GL) was used. The pulse generator uses a coaxial reed switch to discharge a charge line for a very fast rise time pulse outputs. The model 732 pulse generator provides pulses of less than 50 picoseconds (ps) rise time, with amplitude from 150 V to greater than 2 KV into any load impedance through a 50 ohm coaxial line. The generator is capable of producing pulses with a minimum width of 750 ps and a maximum of 1 microsecond. This output pulse width is determined by charge line length for rectangular pulses, or by capacitors for 1/e decay pulses.

For the return data we used in this paper, each sample is spaced at 50 picoseconds interval, and 16,000 samples were collected for each collection for a total time duration of 0.8 microsec-

onds at a rate of approximately 20 Hz. We considered two sets of data from this experiment. Initially, the Barth pulse source was operated at lower amplitude and 35 pulses of clutter signals were obtained at each site but different time. These pulses have been averaged to remove the random noise. Data have been collected from 10 different sites. one collection of transmitted pulse and received backscattering are shown in Fig. 6(a) and (b) respectively. Although pulse-to-pulse variability was noted for collections of received echoes, the fading tendency of different returned signals are the same. These data is referred to as data set I.

Later, additional improvements were made in the measurement procedure, include the improved isolation of transmit and receive antennas, the addition of a log-periodic antenna (Antenna Research Associates LPC-2010-C) as a transmit antenna, and the EMCO ridged waveguide horn (Microwave horn, EMCO 3106). Echoes for data set II were collected using this higher amplitude transmitted pulses. 2 collections at different site with 100 pulsese average have been obtained, one of which is shown in Fig. 6(c). To make them clearer to readers, we provide expanded views of received traces from sample 10,000 to 12,000 in Fig. 7.

4 Statistical Analysis of the Foliage Clutter Data

4.1 Maximum Likelihood Estimation

Using the collected clutter data mentioned above, we apply Maximum Likelihood Estimation (MLE) approach to estimate the parameters of the log-logistic, log-normal, Weibull, and Nakagami models. MLE is often used when the sample data are known and parameters of the underlying probability distribution are to be estimated [32] [33]. It is generalized as follows:

Let $y_1, y_2, \cdots, y_N$ be N independent samples drawn from a random variable $\mathbf{Y}$ with m parameters $\theta_1, \theta_2, \cdots, \theta_m$, where $\theta_i \in \theta$, then the likelihood function expressed as a function of θ conditional on $\mathbf{Y}$ is

$$L_N(\mathbf{Y}|\theta) = \prod_{k=1}^{N} f_{Y|\theta}(y_k|\theta_1, \theta_2, \cdots, \theta_m) \tag{18}$$

The maximum likelihood estimate of θ_1, θ_2, $\cdots$, θ_m is the set of values $\hat{\theta}_1$, $\hat{\theta}_2$, $\cdots$, $\hat{\theta}_m$ that maximize the likelihood function $L_N(\mathbf{Y}|\theta)$.

As the logarithmic function is monotonically increasing, maximizing $L_N(\mathbf{Y}|\theta)$ is equivalent to maximizing $\ln(L_N(\mathbf{Y}|\theta))$. Hence, it can be shown that a necessary but not sufficient condition to obtain the ML estimate $\hat{\theta}$ is to solve the likelihood equation

$$\frac{\partial}{\partial\theta}\ln(L_N(\mathbf{Y}|\theta)) = 0 \tag{19}$$

Note that the amplitude of foliage clutter faded with the increase of sample time. Even At the same sample, it varies for different collections. In order to better analyze its randomness, we studied each collection. Using the collected clutter radar mentioned above, we apply MLE to obtain $\hat{\mu}$ and $\hat{\sigma}$ for log-logistic, $\hat{\mu}$ and $\hat{\sigma}$ for the log-normal, $\hat{a}$ and $\hat{b}$ for the Weibull, and $\hat{\mu}$ and $\hat{\omega}$ for the Nakagami. The estimation results for data set I are listed in table 1. We also explore the standard deviation (STD) error of each parameter. These descriptions are shown in table 1 in the form of ε_x, where x denotes different parameter for each model. We also calculate the average values of estimated parameters and their STD errors in table 2.

From table 1 and 2 we can see STD error for log-logistic and log-normal parameters are less than 0.02 and their estimated parameters vary little from data to data compared to Weibull and Nakagami. It is obvious that log-logistic model provides the smallest STD error for all the 10 collections compared to log-normal. It is obvious that apply both Weibull and Nakagami models, accurate shape parameter estimation can be achieved but the result of scale parameter estimations are not acceptable.

The estimation results for data set II are shown in table 3. Due to the improvement on this set of signal, STD error for log-logistic and log-normal parameters have been reduced compared those of data set I. However, for Weibull and Nakagami, it is a different case, which implies log-logistic and log-normal are much more accurate to model foliage clutter.

In the view of error on parameter estimation, log-logistic model fits the collected data best compared to log-normal, Weibull, and Nakagami. Log-normal is also acceptable.

4.2 Goodness-of-fit in curve and RMSE

We may also observe the extend to which the PDF curve of the statistic model matches that of clutter data by calculating the averaged root mean square error (RMSE) for each data set. Let i (i=$1, 2, \cdots, n$) be the sample index of clutter amplitude, c_i is the corresponding PDF value whereas $\hat{c}_i$ is the PDF value of the statistical model with estimated parameters by means of MSE. The RMSE is obtained through

$$\texttt{RMSE} = \frac{1}{k}\sum_{k}\sqrt{\frac{1}{n}\sum_{i=1}^{n}(c_i - \hat{c}_i)^2} \tag{20}$$

Here we apply n=100 for each model and k is the number of data collections for each set.

In Fig. 8 and 9, we use one collection from data set I and II respectively to illustrate the goodness-of-fit in curve. Also, we calculate the averaged RMSE of each model for both collected data set I and II. The PDF of absolute amplitude of one collection of clutter data is presented by means of histogram bars. In Fig. 8, it can be seen obviously that log-logistic model with MLE parameters provides the best goodness-of-fit compared to the other models, since it provides the most suitable kurtosis, slope and tail. As for the maximum PDF value, the log-logistic is about 1×10^{-3}, while that of other models are over 1.2×10^{-3}. For the slope part, which connects the kurtosis and the tail and which is in the range from 0.1×10^4 to 0.5×10^4 in view of x axes, the log-logistic provides the smallest skewness whereas Nakagami provides the largest. Examination of the tails show that log-logistic and log-normal provide very similar-valued tails, while tails of the Weibull and the Nakagami are lager than the collected data. Meanwhile, we obtain that $\texttt{RMSE}_{log-logistic} = 2.5425 \times 10^{-5}$, $\texttt{RMSE}_{log-normal} = 3.2704 \times 10^{-5}$, $\texttt{RMSE}_{Weibull} = 3.7234 \times 10^{-5}$, $\texttt{RMSE}_{Nakagami} = 5.4326 \times 10^{-5}$. This also shows that the log-logistic model is more accurate than the other three models.

Similarly, in Fig. 9 histogram bars denote the PDF of the absolute amplitude of one collection of clutter data from set II. Compared to Fig 8, the log-logistic and the log-normal provide more similar extend of goodness-of-fit. Weibull is worse since it cannot fit well

in either kurtosis or tail, while Nakagami is the worst and unacceptable. Also, we obtain $\texttt{RMSE}_{log-logistic} = 2.739 \times 10^{-5}$, $\texttt{RMSE}_{log-normal} = 3.1866 \times 10^{-5}$, $\texttt{RMSE}_{Weibull} = 3.6361 \times 10^{-5}$, $\texttt{RMSE}_{Nakagami} = 4.4045 \times 10^{-5}$. This illustrates that for clutter backscattering data set II, the log-logistic model still fits the best.

5 Target Detection Performance

As we have mentioned previously, one of the primary goal to be carried out by a radar is target detection. On a basis of the clutter model that have been just studied, we may apply a special case of the Bayesian criterion named Neyman-Person criterion to analyze the target detection performance in the foliage environment.

If the received sample signal to be tested is R, then the two hypotheses are shown as follows:

$$\begin{aligned} H_0 &: R = C + n \\ H_1 &: R = S + C + n \end{aligned} \tag{21}$$

where C and n represent the random variable of clutter and noise respectively. C follows log-logistic model with both parameters μ and σ, and n is gaussian noise with zero mean and variance ν^2. S is the target signal, which assumes to be a constant for simplicity.

Therefore $f(R|H_0)$ and $f(R|H_1)$ mean:

$f(R|H_0) =$ *PDF of R given that a target was not present*

$f(R|H_1) =$ *PDF of R given that a target was present*

They can be denoted as follows:

$$f(R|H_0) = \int_0^\infty \frac{e^{\frac{lnr-\mu}{\sigma}}}{\sigma r(1+e^{\frac{lnr-\mu}{\sigma}})^2} \times \frac{1}{\sqrt{2\pi}\nu} e^{-\frac{(R-r)^2}{2\nu^2}} dr \tag{22}$$

$$f(R|H_1) = \int_0^\infty \frac{e^{\frac{ln(r-s)-\mu}{\sigma}}}{\sigma(r-s)(1+e^{\frac{ln(r-s)-\mu}{\sigma}})^2} \times \frac{1}{\sqrt{2\pi}\nu} e^{-\frac{(R-s-r)^2}{2\nu^2}} dr \tag{23}$$

If the probability that a target was not present is $P(H_0)$ whereas that of a target was present is $P(H_1)$, then PDF of R is

$$f(R) = P(H_0)f(R|H_0) + P(H_1)f(R|H_1) \tag{24}$$

To decide whether there is a target or not, Neyman-Pearson detection rule is shown as

$$\frac{f(R|H_0)}{f(R|H_1)} \underset{H_1}{\overset{H_0}{\gtrless}} \frac{P(H_1)}{P(H_0)} \tag{25}$$

In case of $P(H_1) = P(H_0)$, (26) is simplified as

$$f(R|H_0) \underset{H_1}{\overset{H_0}{\gtrless}} f(R|H_1) \tag{26}$$

which actually is

$$\frac{e^{[\frac{s^2-2s(R-r)}{2\nu^2}+\frac{ln(\frac{r}{r-s})}{\sigma}]}}{\frac{r}{r-s}[\frac{1+e^{\frac{lnr-\mu}{\sigma}}}{1+e^{\frac{ln(r-s)-\mu}{\sigma}}}]^2} \underset{H_1}{\overset{H_0}{\gtrless}} 1 \tag{27}$$

It is easy to obtain the decision threshold T in terms of the above function

$$T = -\frac{\nu^2}{s} ln[\frac{1+e^{\frac{lnr-\mu}{\sigma}}}{1+e^{\frac{ln(r-s)-\mu}{\sigma}}}]^2 + \frac{\nu^2[ln(\frac{r}{r-s})-\sigma]}{s\sigma} + \frac{s}{2} + r \tag{28}$$

Under hypothesis H_0, a false alarm occurs anytime $R > T$, therefore the probability of false alarm is

$$\begin{aligned} P_{FA} &= \int_T^\infty f(R|H_0)dR \\ &= \frac{1}{\sqrt{2\pi}\sigma\nu}\int_T^\infty \int_0^\infty \frac{e^{[-\frac{(R-r)^2}{2\nu^2}+\frac{lnr-\mu}{\sigma}]}}{(1+e^{\frac{lnr-\mu}{\sigma}})^2 r} drdR \end{aligned} \tag{29}$$

Similarly, Under hypothesis H_1, when $R > T$, the target is detectable. Therefore the probability of detection is

$$\begin{aligned} P_D &= \int_T^\infty f(R|H_1)dR \\ &= \frac{1}{\sqrt{2\pi}\sigma\nu}\int_T^\infty \int_0^\infty \frac{e^{[-\frac{(R-r-s)^2}{2\nu^2}+\frac{ln(r-s)-\mu}{\sigma}]}}{(1+e^{\frac{ln(r-s)-\mu}{\sigma}})^2 (r-s)} drdR \end{aligned} \tag{30}$$

6 Conclusion

On a basis of 2 sets of foliage clutter data using UWB radar, we show that it is more accurate to describe the amplitude of foliage clutter using log-logistic statistic model rather than log-normal, Weibull, or Nakagami. Log-normal is also acceptable. The goodness-of-fit for Weibull

is worse whereas that of Nakagami is the worst. Our contribution is not only the new proposal on the foliage clutter model with estimated concrete parameters, but also provide the criteria and approaches based on which the statistical analysis is deduced. Further, the theoretical study about the probability of detection as well as the probability of false alarm at the presence of foliage clutter and white gaussian noise is discussed. Future research will investigate the characteristics of targets and the design of radar receiver to better achieve the target detection, tracking and imaging.

Acknowledgement

This work was supported in part by Office of Naval Research (ONR) under Grant N00014-07-1-0395, N00014-07-1-1024.

References

[1] ONR BAA 07-009, "Electronic Warfare Discovery and Invention (D&I)," http://www.onr.navy.mil/02/baa/.

[2] Imhoff, M. L., "A theoretical analysis of the effect of forest structure on SAR backscatter and the remote sensing of biomass," *IEEE Trans. Geosci. Remote Sensing*, vol. 33, pp. 341C352, Mar. 1995.

[3] Skolnik, M. I.,*Introduction to Radar Systems*, 3rd ed, New York, McGraw Hill, 2001.

[4] Posner, F. L., *IEEE Trans. on Aerosp. Electron. Syst.*, vol. 38, no. 1, pp. 58-73, Jan 2002.

[5] David, A. Shnidman, "Generalized Radar Clutter Model", *IEEE Trans. on Aerosp. Electron. Syst*, vol. 35, no. 3, July 1999.

[6] Hennessey, G. and Leung, H., "Sea-Clutter Modeling Using a Radial-Basis-Function Neural Network", *IEEE Journal of Oceanic Engineering*, vol. 26, no. 3, pp. 358-372, July, 2001.

[7] Xie, N. and Leung, H., "A Multiple-Model Prediction Approach for Sea Clutter Modeling", *IEEE Trans. on Geoscience and Remote Sensing*, vol. 41, no. 6, pp. 1491-1502, June, 2003.

[8] Fleischman, J. G., Ayasli, S., Adams, E. M., and Gosselin, D. R., Foliage penetration experiment: Part I: Foliage attenuation and backscatter analysis of SAR imagery, *IEEE Trans. on Aerosp. Electron. Syst.*, 32, 1, part 1 of 3 (1996), 134-144.

[9] Mccorkle, J. W., "Early results from the ARL UWB Foliage attenuation (FOPEN) SAR", Presented at the SPIE International Symposium on Optical Engineering and Photonics in Aerospace and Remote Sensing, Conference 1942, Underground and Obscured Object Detection, Apr. 1993.

[10] Sheen, D. R., Malinas, N. P., Kletzli, D. W., Lewis, T. B., and Roman, J. F., "Foliage transmission meassurement using a ground-based ultra-wideband (UWB) (300-1300MHz) SAR system", *IEEE Trans. on Geoscience and Remote Sensing*, 32, 1(1994).

[11] Watts, S., "Radar detection prediction in K-distribution sea clutter and thermal noise", *IEEE Trans. on Aerosp. Electron. Syst.*, AES-23, 1, 1987, pp. 40-45.

[12] Dill, C., "Foliage Penetration (Phase II) Field Test: Narrowband versus Wideband Foliage Penetration," Final Report of Contract Number F41624-03-D-7001/04, July 2005 to Feb 2006.

[13] Gupta, R. C., Akman, O. and Lvin, S., "A Study of Log-Logistic Model in Survival Analysis", *Biometrical Journal*, 41, pp. 431-443, 1999

[14] Singh, V. P., Guo, H. and Yu, F. X., "Parameter estimation for 3-parameter log-logistic distribution (LLD3) by Pome", *Parameter estimation for 3-parameter log-logistic distribution (LLD3) by Pome*, vol.7, no.3, pp. 163-177, 2005

[15] Burr, I. W.,"Cumulative frequency functions." *Ann. Math. Statist.*, 13, 215-232, 1942.

[16] Mielke, P. W. and Johnson. E. S.,"Three-parameter kappa distribution maximum likelihood estimates and likelihood ratio tests," *Monthly Weather Rev.*, 101, 701-709, 1973.

[17] Lee, K. S., Sadeghipour, J. and Dracup, J. A., "An approach for frequency analysis of multiyear dought duration.", *Wat. Resour. Res.* 22(5), 655-662, 1986.

[18] Shoukri, M. M., Mian, I. U. H. and Tracy, D. S., " Sampling properties of estimators of the log-logistic distribution with application to Canadian precipitation data.", *Can. J. Statist.* 16(3), 223-236, 1988.

[19] Narda, N. K. and Malik, R. K., "Dynamic model of root growth and water uptake in wheat.", *Indian J. Agric. Engng* 3(3&4), 147-155, 1993.

[20] Warden, M., "An experimental study of some clutter characteristics.", in *AGARD Conf. Proc. 66-Advanced Radar Systems*, May 1970.

[21] Trunk, G. and George, S., "Detection of targets in non-Gaussian sea clutter", *IEEE Trans. on Aerosp. Electron. Syst.*, vol. AES-6, Sept. 1970

[22] Schleher, D. C., "Radar Detection in Weibull Clutter", *IEEE Trans. on Aerosp. Electron. Syst.*, vol. AES-12, No. 6, Nov. 1976

[23] Limpert, E., Stahel, W. and Abbt, M., "Log-normal Distributions across the Sciences: Keys and Clues.", *BioScience*, 51 (5), pp. 341C352, 2001.

[24] Weibull, W. "A statistical distribution function of wide applicability", J. Appl. Mech.-Trans. ASME 18(3), 293-297, 1951.

[25] Boothe, R. R., "The Weibull distribution applied to the ground clutter backscatter coefficient.", U.S. Army Missile Command Report No. RE-TR-69-15, June, 1969.

[26] Sekine, M., *et al*, "Weibull distributed ground clutter.", *IEEE Trans. on Aerosp. Electron. Syst.*, AES-17 pp. 596-598, July 1981.

[27] Fay, F. A., Clarke, J. and Peters, R. S, "Weibull distributed applied to sea clutter", *Radar 77, IEE Conf. Publ.*, 155, pp. 101-104, 1977.

[28] Sekine, M., *et al*, "Weibull distributed sea clutter.", *IEE Proceedings, Part F - Communications, Radar and Signal Processing*, vol. 130, no. 5, Aug. 1983.

[29] Sekine, M., *et al*, "On Weibull distributed weather clutter.", *IEEE Trans. on Aerosp. Electron. Syst.*, AES-15, pp. 824-830.

[30] Tsihrintzis, G. A. and Nikias, C. L., "Evaluation of fractional lower-order statistics-based detection algorithms on real sea-clutter data.", *IEE Proc-Radar, Sonar Navig.*, vol. 144, no.1, Feb., pp. 29-37, 1997.

[31] M. Nakagami, "The m-Distribution, a general formula of intensity of rapid fading". In W. G. Hoffman, editor, *Statistical Methods in Radio Wave Propagation: Proceedings of a Symposium held at the University of California*, pp 3-36. Permagon Press, 1960.

[32] Devore, Probability and Statistics for Engineering and the Sciences. Monterey, CA: BrooksiCole, 1982.

[33] Barkat, M., Signal deteciton and estimation, 2nd, London: Artech house, 2005.

List of Tables

Table 1: Estimated Parameters for Data Set I

PDF	Log-Logistic	Log-normal	Weibull	Nakagami
data 1	$\hat{\mu} = 7.24161$ $\hat{\sigma} = 1.06483$ $\varepsilon_\mu = 0.0141212$ $\varepsilon_\sigma = 0.00724181$	$\hat{\mu} = 7.0455$ $\hat{\sigma} = 2.20761$ $\varepsilon_\mu = 0.0174527$ $\varepsilon_\sigma = 0.0123415$	$\hat{a} = 2975.33$ $\hat{b} = 0.594979$ $\varepsilon_a = 41.6157$ $\varepsilon_b = 0.00356925$	$\hat{\mu} = 0.177062$ $\hat{\omega} = 9.09663e+007$ $\varepsilon_\mu = 0.00150615$ $\varepsilon_\omega = 1.70907e+006$
data 2	$\hat{\mu} = 6.9716$ $\hat{\sigma} = 1.2126$ $\varepsilon_\mu = 0.014747$ $\varepsilon_\sigma = 0.00773723$	$\hat{\mu} = 6.72573$ $\hat{\sigma} = 2.33617$ $\varepsilon_\mu = 0.0184691$ $\varepsilon_\sigma = 0.0130602$	$\hat{a} = 2285.13$ $\hat{b} = 0.563747$ $\varepsilon_a = 33.7127$ $\varepsilon_b = 0.00337485$	$\hat{\mu} = 0.162375$ $\hat{\omega} = 7.4776e+007$ $\varepsilon_\mu = 0.00137422$ $\varepsilon_\omega = 1.46679e+006$
data 3	$\hat{\mu} = 7.00554$ $\hat{\sigma} = 1.10741$ $\varepsilon_\mu = 0.0145728$ $\varepsilon_\sigma = 0.0076303$	$\hat{\mu} = 6.76262$ $\hat{\sigma} = 2.31258$ $\varepsilon_\mu = 0.0182825$ $\varepsilon_\sigma = 0.0129283$	$\hat{a} = 2341.52$ $\hat{b} = 0.57073$ $\varepsilon_a = 34.1207$ $\varepsilon_b = 0.00341448$	$\hat{\mu} = 0.164695$ $\hat{\omega} = 7.46366e+007$ $\varepsilon_\mu = 0.001395$ $\varepsilon_\omega = 1.45459e+006$
data 4	$\hat{\mu} = 7.03055$ $\hat{\sigma} = 1.07858$ $\varepsilon_\mu = 0.0142027$ $\varepsilon_\sigma = 0.00741556$	$\hat{\mu} = 6.80711$ $\hat{\sigma} = 2.25973$ $\varepsilon_\mu = 0.0178647$ $\varepsilon_\sigma = 0.0126329$	$\hat{a} = 2395.85$ $\hat{b} = 0.579381$ $\varepsilon_a = 34.4066$ $\varepsilon_b = 0.00345156$	$\hat{\mu} = 0.167391$ $\hat{\omega} = 7.4926e+007$ $\varepsilon_\mu = 0.0014916$ $\varepsilon_\omega = 1.44727e+006$
data 5	$\hat{\mu} = 7.16226$ $\hat{\sigma} = 1.10132$ $\varepsilon_\mu = 0.014605$ $\varepsilon_\sigma = 0.00750067$	$\hat{\mu} = 6.95712$ $\hat{\sigma} = 2.26592$ $\varepsilon_\mu = 0.0179137$ $\varepsilon_\sigma = 0.0126675$	$\hat{a} = 2806.76$ $\hat{b} = 0.577823$ $\varepsilon_a = 40.4226$ $\varepsilon_b = 0.00347389$	$\hat{\mu} = 0.17112$ $\hat{\omega} = 9.03298e+007$ $\varepsilon_\mu = 0.00145265$ $\varepsilon_\omega = 1.72749e+006$
data 6	$\hat{\mu} = 7.01527$ $\hat{\sigma} = 1.10123$ $\varepsilon_\mu = 0.0144902$ $\varepsilon_\sigma = 0.00758568$	$\hat{\mu} = 6.77515$ $\hat{\sigma} = 2.30286$ $\varepsilon_\mu = 0.0182057$ $\varepsilon_\sigma = 0.012874$	$\hat{a} = 2360.33$ $\hat{b} = 0.572749$ $\varepsilon_a = 34.2753$ $\varepsilon_b = 0.00342376$	$\hat{\mu} = 0.165292$ $\hat{\omega} = 7.50824e+007$ $\varepsilon_\mu = 0.00140035$ $\varepsilon_\omega = 1.46145e+006$
data 7	$\hat{\mu} = 7.14523$ $\hat{\sigma} = 1.09486$ $\varepsilon_\mu = 0.0145132$ $\varepsilon_\sigma = 0.00745994$	$\hat{\mu} = 6.94201$ $\hat{\sigma} = 2.25621$ $\varepsilon_\mu = 0.0178369$ $\varepsilon_\sigma = 0.0126132$	$\hat{a} = 2753.69$ $\hat{b} = 0.578948$ $\varepsilon_a = 39.585$ $\varepsilon_b = 0.00347442$	$\hat{\mu} = 0.170964$ $\hat{\omega} = 8.80474e+007$ $\varepsilon_\mu = 0.00145125$ $\varepsilon_\omega = 1.68382e+006$
data 8	$\hat{\mu} = 6.95411$ $\hat{\sigma} = 1.11486$ $\varepsilon_\mu = 0.0146774$ $\varepsilon_\sigma = 0.00768003$	$\hat{\mu} = 6.71591$ $\hat{\sigma} = 2.31898$ $\varepsilon_\mu = 0.0183331$ $\varepsilon_\sigma = 0.0129641$	$\hat{a} = 2250.66$ $\hat{b} = 0.564989$ $\varepsilon_a = 33.1387$ $\varepsilon_b = 0.0033763$	$\hat{\mu} = 0.162448$ $\hat{\omega} = 7.31436e+007$ $\varepsilon_\mu = 0.00137488$ $\varepsilon_\omega = 1.4338e+006$
data 9	$\hat{\mu} = 7.18561$ $\hat{\sigma} = 1.09854$ $\varepsilon_\mu = 0.0145483$ $\varepsilon_\sigma = 0.00749265$	$\hat{\mu} = 6.9715$ $\hat{\sigma} = 2.27088$ $\varepsilon_\mu = 0.0179529$ $\varepsilon_\sigma = 0.0126952$	$\hat{a} = 2840.72$ $\hat{b} = 0.581219$ $\varepsilon_a = 40.6593$ $\varepsilon_b = 0.0034984$	$\hat{\mu} = 0.172324$ $\hat{\omega} = 8.97304e+007$ $\varepsilon_\mu = 0.00146348$ $\varepsilon_\omega = 1.70923e+006$
data 10	$\hat{\mu} = 7.192$ $\hat{\sigma} = 1.0866$ $\varepsilon_\mu = 0.0144166$ $\varepsilon_\sigma = 0.0073916$	$\hat{\mu} = 6.99196$ $\hat{\sigma} = 2.23975$ $\varepsilon_\mu = 0.0177067$ $\varepsilon_\sigma = 0.0125211$	$\hat{a} = 2869.65$ $\hat{b} = 0.584803$ $\varepsilon_a = 40.837$ $\varepsilon_b = 0.00351294$	$\hat{\mu} = 0.173572$ $\hat{\omega} = 9.01631e+007$ $\varepsilon_\mu = 0.0014747$ $\varepsilon_\omega = 1.71142e+006$

Table 2: Averaged Estimated Parameters for Data Set I

PDF	Log-Logistic	Log-normal	Weibull	Nakagami
average	$\hat{\mu} = 7.0904$ $\hat{\sigma} = 1.1061$ $\varepsilon_\mu = 0.0145$ $\varepsilon_\sigma = 0.0075$	$\hat{\mu} = 6.8695$ $\hat{\sigma} = 2.2771$ $\varepsilon_\mu = 0.0180$ $\varepsilon_\sigma = 0.0127$	$\hat{a} = 2588$ $\hat{b} = 0.5769$ $\varepsilon_a = 37.4316$ $\varepsilon_b = 0.0035$	$\hat{\mu} = 0.1687$ $\hat{\omega} = 8.218e + 007$ $\varepsilon_\mu = 0.0014$ $\varepsilon_\omega = 1.4905e + 006$

Table 3: Estimated and Averaged Parameters for Data Set II

PDF	Log-Logistic	Log-normal	Weibull	Nakagami
data 1	$\hat{\mu} = 7.76868$ $\hat{\sigma} = 0.786511$ $\varepsilon_\mu = 0.0107792$ $\varepsilon_\sigma = 0.00521601$	$\hat{\mu} = 7.79566$ $\hat{\sigma} = 1.41771$ $\varepsilon_\mu = 0.011208$ $\varepsilon_\sigma = 0.00792559$	$\hat{a} = 4901.07$ $\hat{b} = 0.743223$ $\varepsilon_a = 55.3011$ $\varepsilon_b = 0.00434465$	$\hat{\mu} = 0.239587$ $\hat{\omega} = 1.16839e + 008$ $\varepsilon_\mu = 0.00207912$ $\varepsilon_\omega = 1.88719e + 006$
data 2	$\hat{\mu} = 7.78096$ $\hat{\sigma} = 0.787426$ $\varepsilon_\mu = 0.0107917$ $\varepsilon_\sigma = 0.0052213$	$\hat{\mu} = 7.8046$ $\hat{\sigma} = 1.41855$ $\varepsilon_\mu = 0.0112147$ $\varepsilon_\sigma = 0.00793033$	$\hat{a} = 4942.48$ $\hat{b} = 0.745233$ $\varepsilon_a = 55.6114$ $\varepsilon_b = 0.0043612$	$\hat{\mu} = 0.240593$ $\hat{\omega} = 1.17237e + 008$ $\varepsilon_\mu = 0.00208848$ $\varepsilon_\omega = 1.88953e + 006$
average	$\hat{\mu} = 7.7748$ $\hat{\sigma} = 0.7870$ $\varepsilon_\mu = 0.0108$ $\varepsilon_\sigma = 0.0052$	$\hat{\mu} = 7.7881$ $\hat{\sigma} = 1.4181$ $\varepsilon_\mu = 0.0112$ $\varepsilon_\sigma = 0.0079$	$\hat{a} = 4921.8$ $\hat{b} = 0.7442$ $\varepsilon_a = 55.4565$ $\varepsilon_b = 0.0044$	$\hat{\mu} = 0.2401$ $\hat{\omega} = 1.1704 + 008$ $\varepsilon_\mu = 0.0021$ $\varepsilon_\omega = 1.8884 + 006$

List of Figures

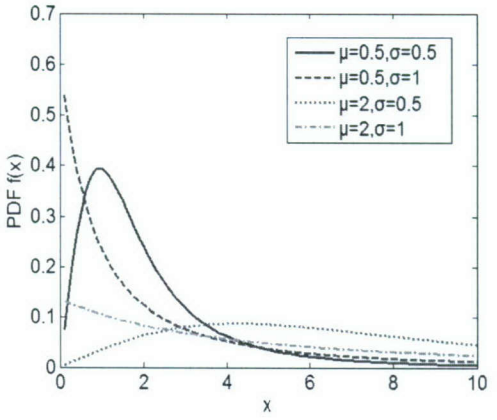


Figure 1: Log-logistic distribution PDF for $\mu = 0.5$ and $\sigma = 0.5$, $\mu = 0.5$ and $\sigma = 1$, $\mu = 2$ and $\sigma = 0.5$, $\mu = 2$ and $\sigma = 1$

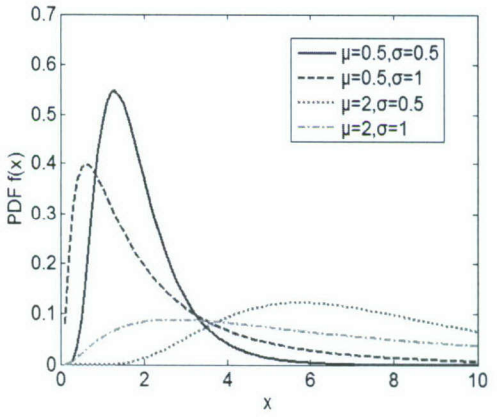


Figure 2: Log-normal distribution PDF for $\mu = 0.5$ and $\sigma = 0.5$, $\mu = 0.5$ and $\sigma = 1$, $\mu = 2$ and $\sigma = 0.5$, $\mu = 2$ and $\sigma = 1$

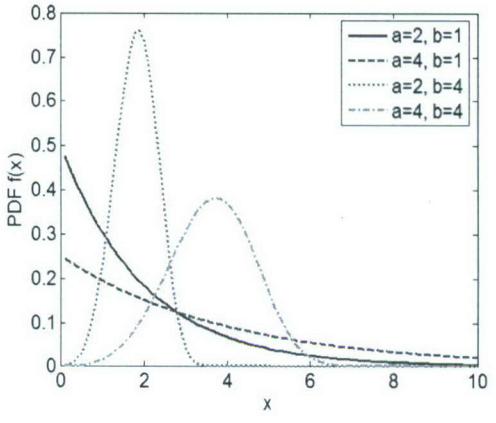


Figure 3: Weibull distribution PDF for $a = 2$ and $b = 1$, $a = 4$ and $b = 1$, $a = 2$ and $b = 4$, $a = 4$ and $b = 4$

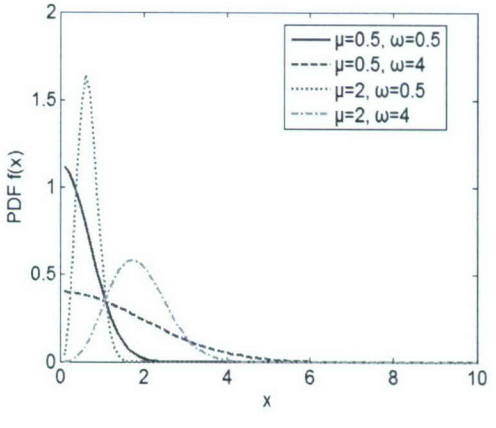
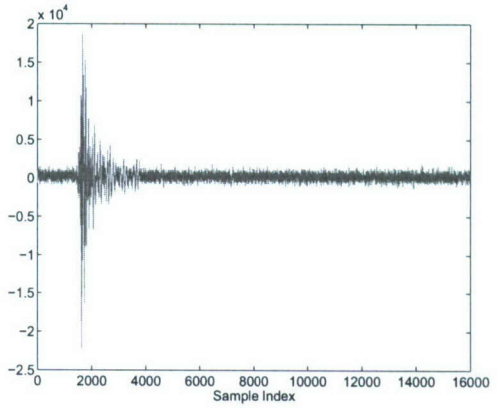
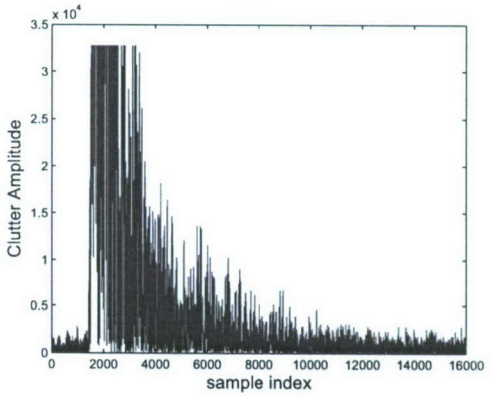


Figure 4: Nakagami distribution PDF for $\mu = 0.5$ and $\omega = 0.5$, $\mu = 0.5$ and $\omega = 4$, $\mu = 2$ and $\omega = 0.5$, $\mu = 2$ and $\omega = 4$

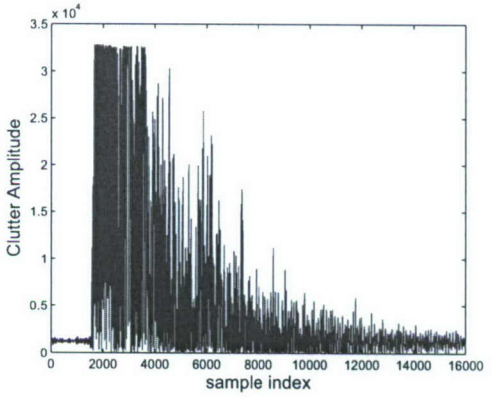
Figure 5: This figure shows the lift with the experiment. The antennas are at the far end of the lift from the viewer under the roof that was built to shield the equipment from the elements. This picture was taken in September with the foliage largely still present. The cables coming from the lift are a ground cable to an earth ground and one of 4 tethers used in windy conditions.



(a)



(b)



(c)

Figure 6: Clutter data (a)transmitted pulse before antenna amplification (b) an example of received echoes from data set I, and (c) an example of received echoes from data set II

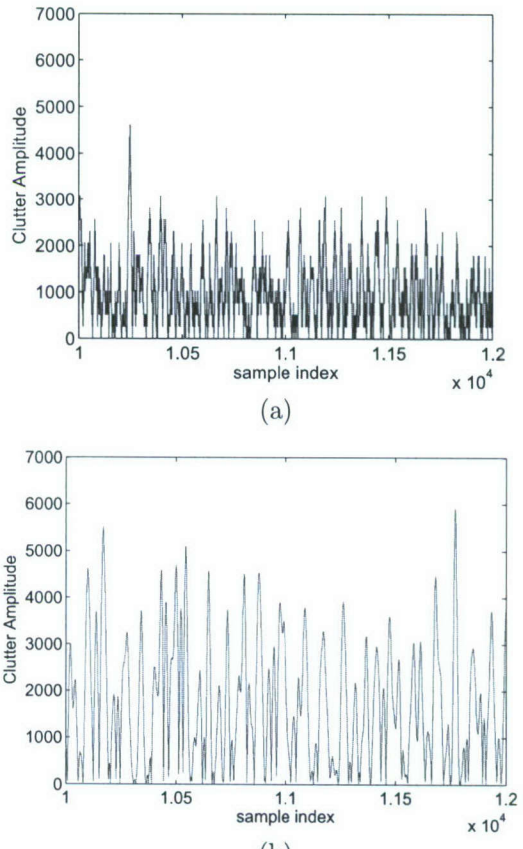
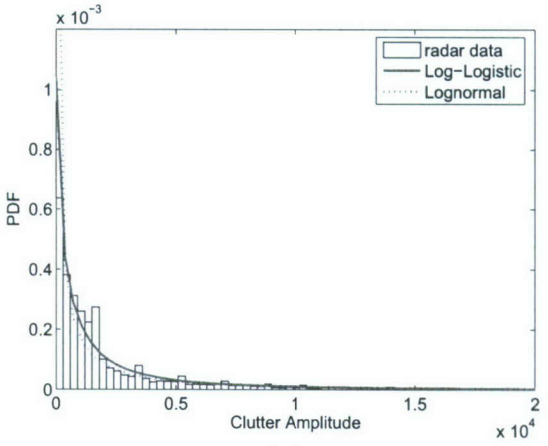
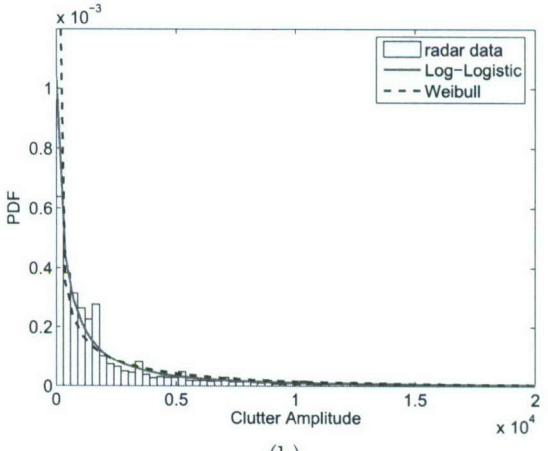


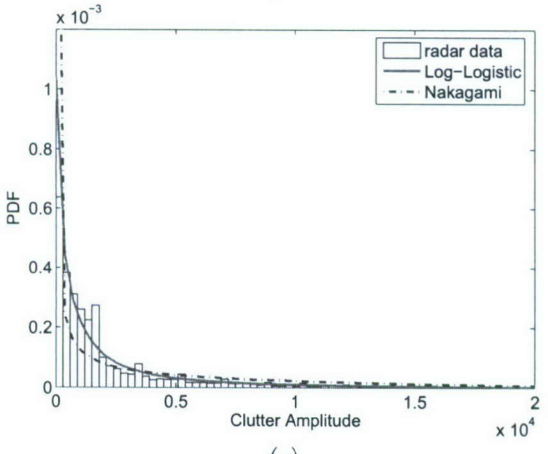
Figure 7: Expanded view from clutter samples 10,000 to 12,000 (a) from data set I, and (b) from data set II



(a)

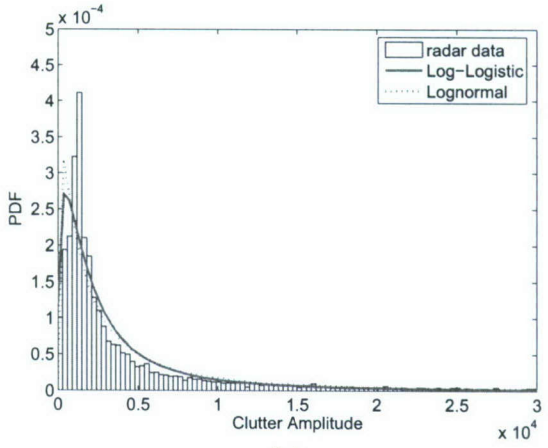


(b)

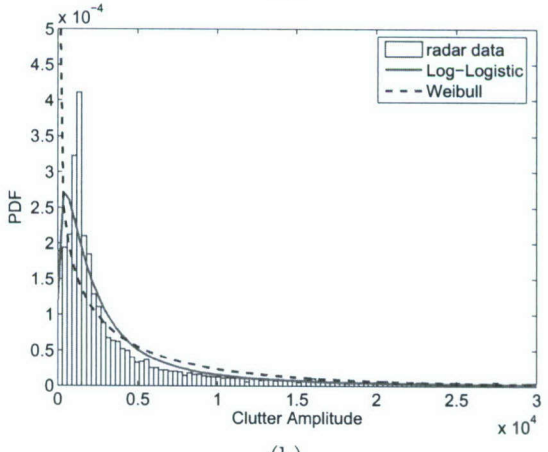


(c)

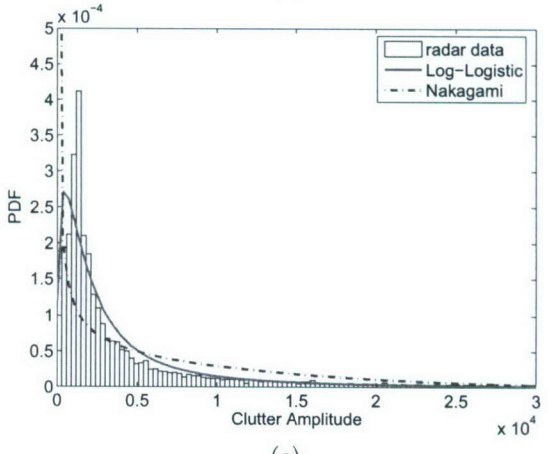
Figure 8: Clutter model comparison from data set I (a) log-logistic vs. log-normal, and (b) log-logistic vs. Weibull (c) log-logistic vs. Nakagami. $\texttt{RMSE}_{log-logistic} = 2.5425 \times 10^{-5}$, $\texttt{RMSE}_{log-normal} = 3.2704 \times 10^{-5}$, $\texttt{RMSE}_{Weibull} = 3.7234 \times 10^{-5}$, $\texttt{RMSE}_{Nakagami} = 5.4326 \times 10^{-5}$.



(a)



(b)



(c)

Figure 9: Clutter model comparison from data set II (a) log-Logistic vs. log-normal, and (b) log-logistic vs. Weibull (c) log-logistic vs. Nakagami. $\texttt{RMSE}_{log-logistic} = 2.739 \times 10^{-5}$, $\texttt{RMSE}_{log-normal} = 3.1866 \times 10^{-5}$, $\texttt{RMSE}_{Weibull} = 3.6361 \times 10^{-5}$, $\texttt{RMSE}_{Nakagami} = 4.4045 \times 10^{-5}$.

Target Detection in Foliage Using Short-Time Fourier Transform and UWB Radar Sensor Networks

Jing Liang, Qilian Liang
Department of Electrical Engineering
University of Texas at Arlington
Arlington, TX 76019-0016, USA
Email: jliang@wcn.uta.edu, liang@uta.edu

Abstract— **In this paper, we study sense-through-foliage target detection. When radar echoes are in good quality, the detection of target can be achieved by applying short time Fourier transform (STFT) to the received UWB radar waveform. We compared our approach in case of no target as well as with target against the scheme in which 2-D image was created via adding voltages with the appropriate time offset. Results show that our approach can detect target more easily. When radar echoes are in poor condition and single radar is unable to carry out the detection, we employ both Radar Sensor Networks (RSN) and RAKE structure to combine the echoes from different radar members and finally detect the target.**

I. Introduction

Detection and identification of military equipment in a strong clutter background, such as foliage, soil cover or building leads has been a long-standing subject of intensive study. It is believed that solving the target detection through foliage will significantly benefit sense-through-wall and many other subsurface sensing problems. However, to this date, the detection of foliage-covered military targets, such as artillery, tanks, trucks and other weapons with the required probability of detection and false alarm still remains a challenging issue. This is due to the following facts:

1) Given certain low radar cross section(RCS), scattering from tree trunk and ground reflectivity may largely overwhelm the returned target signals of interest, thus very high multiple fading may severely corrupt the amplitude and phase of the echoes.
2) Even if the target is stationery, tree leaves and branches are likely to swing in result of gust, which will result in doppler shift of clutter and difficulty of target detection.

Over the past two decades, following 3 types of signals have been mainly studied to examine the performance on target detection in foliage:

1) Traditional sinusoidal waveforms at VHF through UHF bands [1], as the lower the radar frequency, the lower the attenuation and scattering from branches and trees, and thus better penetration through foliage. However, these approaches result in low resolution and low RCS.
2) Millimeter-Wave (MMW) radars are used in [2] [3] and [4]. Results demonstrate the potential for satisfying performance but need further investigation.
3) Relatively low frequency Ultra-wide band (UWB) radars between 100 MHZ and 3 GHz are frequently employed in recent years owning to the characteristics provided by their high resolutions as well as the very good ability of penetration, such as penetrating walls [5] [6]. Despite comparatively short detection range, UWB signal would have advantages over a narrowband signal with limited frequency content.

In this paper, we will apply our expertise in signal processing, data fusion, radar sensor networks (RSN) etc. to achieve effective target detection in foliage using ultra-wideband (UWB) radar.

The remainder of this paper is organized as follows. In Section II, we summarize the measurement and collection of data used in this study. In Section III, we propose a short time Fourier transform (STFT) based approach for through-foliage target detection when the signal quality is good. In Section IV, we propose RSN and RAKE structure for target detection in foliage when the signal quality is poor. We conclude this paper and discuss some future research topics in Section V.

II. Data Measurement and Collection

Our work is based on the foliage penetration field test data from Air Force Research Lab [7]. The foliage penetration measurement effort began in August 2005 and continued through December 2005. Working in August through the fall of 2005, the foliage measured included late summer foliage and fall and early winter foliage. Late summer foliage, because of the limited rainfall, involved foliage with decreased water content. Late fall and winter measurements involved largely defoliated but dense forest.

The foliage experiment was constructed on a seven-ton man lift, which had a total lifting capacity of 450 kg. The limit of the lifting capacity was reached during the experiment as essentially the entire measuring apparatus was placed on the lift. The principle pieces of equipment secured on the lift are: Barth pulser, Tektronix model 7704 B oscilloscope,

dual antenna mounting stand, two antennas, rack system, IBM laptop, HP signal Generator, Custom RF switch and power supply and Weather shield (small hut).

The target is a trihedral reflector with a slant length of 1.5 meters(as shown in Fig. 1). It is 250 feet one-way distance from the antenna. Throughout this work, a Barth pulse source (Barth Electronics, Inc. model 732 GL) was used. The pulse generator uses a coaxial reed switch to discharge a charge line for a very fast rise time pulse outputs. The model 732 pulse generator provides pulses of less than 50 picoseconds (ps) rise time, with amplitude from 150 V to greater than 2 KV into any load impedance through a 50 ohm coaxial line. The generator is capable of producing pulses with a minimum width of 750 ps and a maximum of 1 microsecond. This output pulse width is determined by charge line length for rectangular pulses, or by capacitors for 1/e decay pulses.

Fig. 1. The target (a trihedral reflector) is shown on the stand at 300 feet from the lift.

For the data we used in this paper, each sample is spaced at 50 picosecond interval, and 16,000 samples were collected for each collection for a total time duration of 0.8 microseconds. We considered two sets of data from this experiment. Initially, the Barth pulse source was operated at low amplitude and significant pulse-to-pulse variability was noted for these collections. We refer this set of collections as "poor" signal. The scheme for the target detection in foliage with "poor" signal quality will be presented in Section IV. Later, good signal quality data were collected using higher amplitude pulses and 100 pulses reflected signals were averaged for each collection. The scheme for target detection with "good" signal quality will be presented in Section III.

III. Target Detection with Good Signal Quality: A STFT-based approach

In Fig. 2, we plot two collections with good signal quality, one without a target on range (Fig. 2a) and the other one with a target on range (Fig. 2b and target appears at around sample 14,000). To make it more clear to the readers, we provide expanded views of traces (with target) from sample 13,001 to 15,000 for the above two collections in Figs. 3a and 3b. The target response will be the echo difference between Fig. 3b and Fig. 3a, which is plotted in Fig. 3c. However, in a practical situation we either obtain Fig. 3a (clutter echo without target) or Fig. 3b (target on range) without the knowledge about the presence of a target. The challenge is how can we make target detection only based on Fig. 3b (with target) or Fig. 3a (no target)?

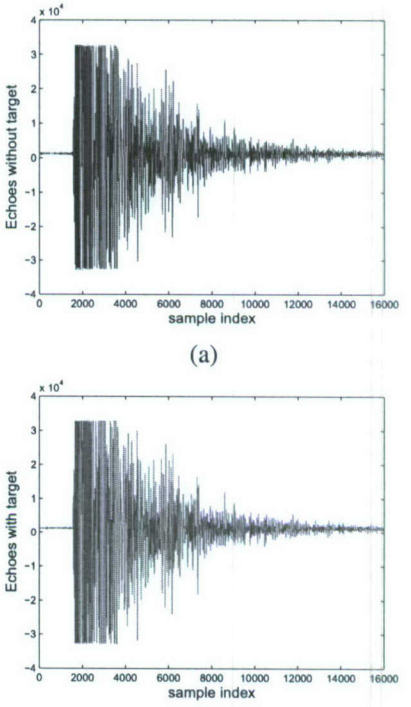
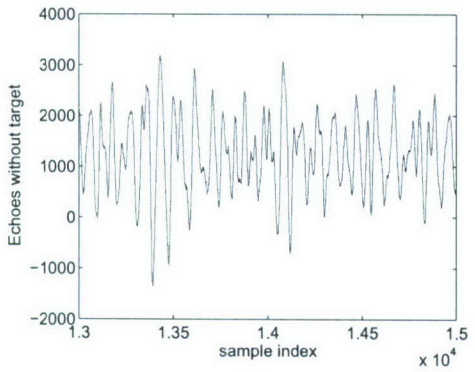


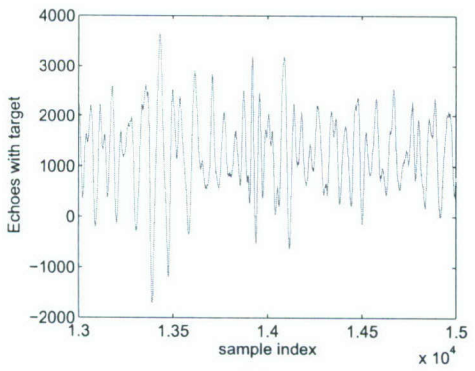
Fig. 2. Measurement with very good signal quality and 100 pulses integration. (a) no target on range (b) with target on range (target appears at around sample 14000)

To solve this problem, a scheme is previously proposed in [8], where 2-D image was created via adding voltages with the appropriate time offset. In Figs. 4(a) and 4(b), we plot the 2-D image created based on the above two data sets (from samples 13,800 to 14,200) using the approach in [8]. However, from these two figures, it's not obvious that which image shows there is a target on range.

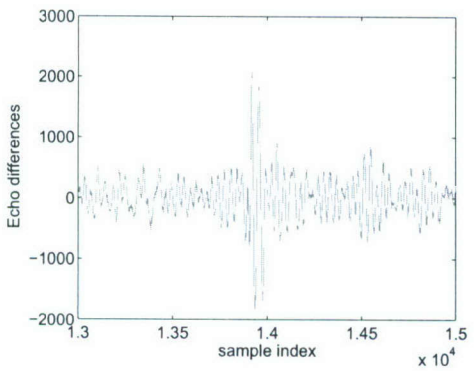
Although Fig. 3a and 3b look quite similar at first sight, after careful observation, it is not difficult to find that the sample strength change more abruptly where target appears (around sample 14000), which implies that echo from target contains more AC values than that without target. This phenomenon inspires the application of short time Fourier transform (STFT), that is using a slide window to determine the sinusoidal frequency and phase content of a signal as it changes over time. This form of the Fourier transform has a great many applications in sonar and radar processing. We will show that STFT-based approach is able to make target detection more intuitively and easily.



(a)



(b)



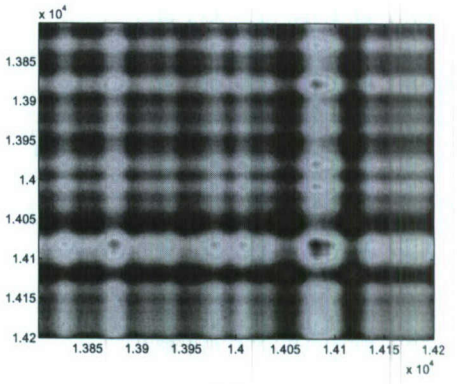
(c)

Fig. 3. Measurement with good signal quality and 100 pulses integration (a) Expanded view of traces (no target) from samples 13001 to 15000 (b) Expanded view of traces (with target) from samples 13001 to 15000 (c) Expanded view of traces difference between with and without target

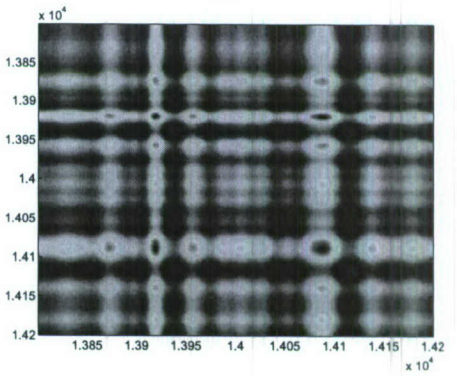
The discrete STFT can be expressed as

$$F(m,w) = \sum_{n=0}^{N-1} r(n)w(n-m)e^{-jwn} \tag{1}$$

where $r(n)$ is radar measurement and $w(n)$ is the window function. We apply rectangular window, with its length $L = 30$ and step size $M = 16$. Then we cumulate the power of AC



(a)



(b)

Fig. 4. 2-D image created via adding voltages with the appropriate time offset (a) no target (b) with target in the field

values ($m \geq 4$)

$$P(m) = \mid \sum_{w=4}^{L-1} F(m,w)^2 \mid \tag{2}$$

We plot the power of AC values $P(m)$ versus time domain sample index in Fig. 5a and 5b for the data sets in Figs. 2a and 2b respectively. We can see that at the samples where there is a target, the curve of the power signal looks like a Gaussian pdf other than chaotic impulses. And thus it is quite straightforward to see that there is no target on range in Fig. 5a.

It's worth mentioning that for better visual inspection, window length and step size may change on a basis of different radar data.

IV. Target Detection with Poor Signal Quality: Radar Sensor Network and Differential-Based Approach

As mentioned in Section II, when the Barth pulse source was operated at low amplitude and the sample values are not obtained based on sufficient pulse response averaging (averaged over 35 pulses for each collection), significant pulse-to-pulse variability was noted and the return signal quality is poor. Fig. 6 illustrate the received echoes in this situation.

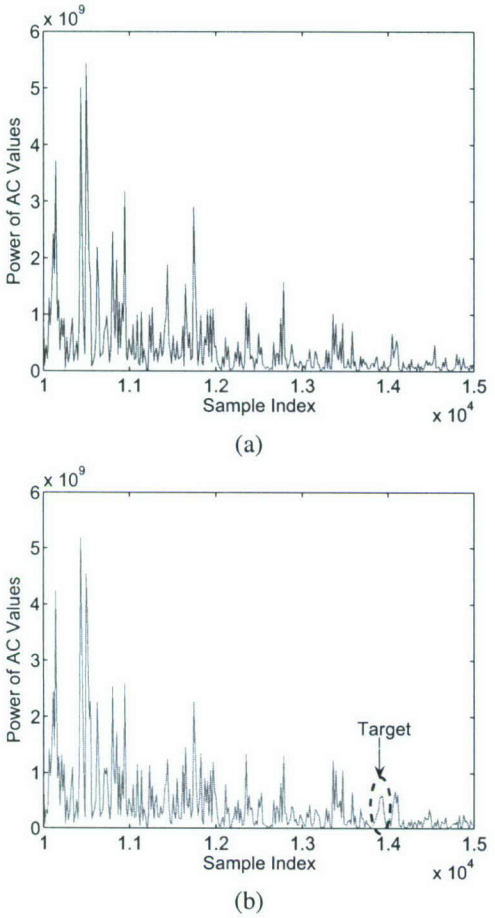


Fig. 5. The power of AC values versus sample index. (a) no target (b) with target in the field

Even with the application of the STFT-based scheme proposed above, we can not tell whether there is a target or not in the range. Since pulse-to-pulse variability exists in the echoes at different time or different site, this motivate us to explore the spatial and time diversity using Radar Sensor Networks (RSN).

In nature, a network of multiple radar sensors can been utilized to combat performance degradation of single radar [10]. These radar sensors are managed by an intelligent clusterhead that combines waveform diversity in order to satisfy the common goals of the network other than each radar operate substantively. As radar sensors are environment dependent [11], it may provide better signal quality if different radars work collaboratively to perform data fusion. For example, consider a system of two radars. When the signal of either radar unfortunately experience a severe fading, if two radars are spaced sufficiently far apart, it is not likely that both of the radars experience deep fade at the same time. By selecting better waveform from the two radar waveforms, the data is less likely to be lost.

In this paper, we assume the radar sensors are synchronized in RSN and we employed RAKE structure to combine received information for RSN. The detailed process is shown in Fig.

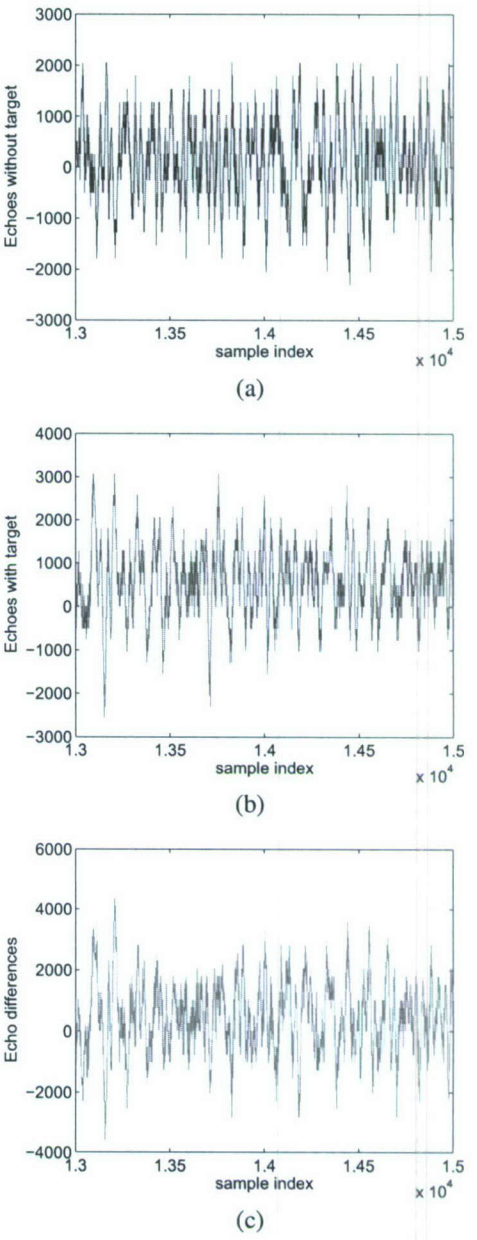


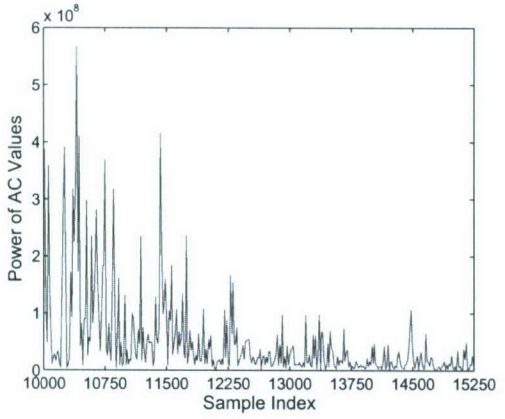
Fig. 6. Measurement with poor signal quality (a) Expanded view of traces (no target) from samples 13001 to 15000 (b) Expanded view of traces (with target) from samples 13001 to 15000 (c) Expanded view of traces difference between with and without target

8. The echo, i.e., RF response by the pulse of each cluster-member radar sensor, will be combined by the clusterhead using a weighted average, and the weight A_i is determined by the power of each echo $x_i(m)$ (m is the sample index),

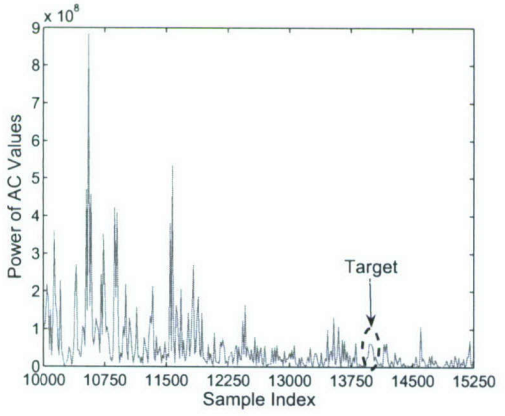
$$A_i = \frac{E_i}{\sum_{i=1}^{n} E_i} \tag{3}$$

and

$$E_i = \text{var}(x_i(m)) + [\text{mean}(x_i(m))]^2 \tag{4}$$



(a)



(b)

Fig. 7. The power of AC values versus sample index. (a) no target (b) with target in the field

As for STFT, we apply window length $L = 25$ and step size $M = 15$. We ran simulations for n = 35 and plot the power of combined signal obtained through STFT based approach in Fig. 7. Compare this figure with Fig. 6a and Fig. 6b, it is quite obvious to see that there is a target around sample 14,000.

V. Conclusion and Future Works

In this paper, we propose a short time Fourier transform (STFT)-based signal processing approach on received UWB Radar waveforms to improve through-foliage target detection. The foliage penetration measurements were taken in Holliston, Massachusetts. When radar echoes are in good quality, the detection of target can be achieved by applying STFT-based technology to single radar. We compared our approach in case of no target as well as with target against the scheme in which 2-D image was created via adding voltages with the appropriate time offset. Results show that our approach can detect target more intuitively. When radar echoes are in poor condition and single radar is unable to carry out detection, we employe both Radar Sensor Networks (RSN) and RAKE structure to combine the echoes from different radar members and finally successfully detect the target. For future works,

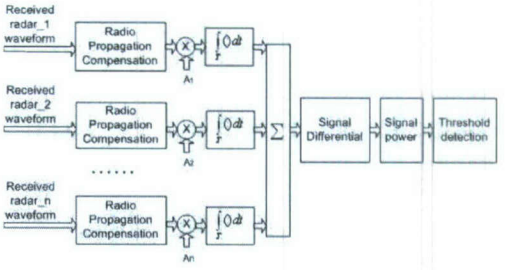


Fig. 8. Block diagram of STFT-based approach and diversity combination in RSN

we will collect more data with different targets and perform automatic target recognition besides target detection.

Acknowledgement

This work was supported in part by Office of Naval Research (ONR) under Grant N00014-07-1-0395, N00014-07-1-1024. The authors would like to thank Dr. Sherwood Samn in AFRL for providing the UWB radar data.

References

[1] S. Ayasli and L. Bessette, " UHF & VHF SAR phenomenology, presented at the Proc. PIERS, Workshop on Advances in Radar Mathods, Baveno, Italy, July 20-22, 1998"

[2] F. K. Schwering, E. J. Violette and R. H. Espeland, "Millimeter-wave propagation in vegetation: Experiments and theory", *IEEE Trans. Geosci. Remote Sensing,* vol.26, pp. 355-367, May 1988

[3] F. T. Ulaby, T. E. Van Deventer, J. R. East, T. F. Haddock and M. E. Coluzzi, "Millimeter-wave bistatic scaterring from ground and vegetation targets" , *IEEE Trans. Geosci. Remote Sensing,* vol.26, pp. 229-243, May 1988

[4] A. Y. Nashashibi, k. Sarabandi, S. Oveisgharan *et al.*, "Millimeter-Wave Measurement of Foliage Attenuation and Ground Reflectivity of Tree Stands at Nadir Incidence", *IEEE Trans. Antennas Propagat.,* vol.52, pp.1211-2004, May, 2004

[5] B. Ferrell "Ultrawideband foliage penetration measurement", in *Proc. IEEE Nation Radar Conf.,* Mar. 29-31, pp.80-84, 1994

[6] X. Xu; R. M. Narayanan, "FOPEN SAR imaging using UWB step-frequency and random noise waveforms ", *IEEE Trans. Aerospace and Electronic Systems*, vol.37, pp.1287-1300, Oct. 2001

[7] C. Dill, "Foliage Penetration (Phase II) Field Test: Narrowband versus Wideband Foliage Penetration," *Final Report of Contract Number F41624-03-D-7001/04*, July 2005 to Feb 2006.

[8] P. Withington , H. Fluhler, and S. Nag, "Enhancing homeland security with advanced UWB sensors," *IEEE Microwave Magazine*, Sept 2003.

[9] Benedetto M. D and Giancola G., *Understanding Ultra Wide Band Radio Fundamentals*, Person Education, 2004

[10] S. Haykin, "Cognitive radar networks" , *2005 1st IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing*, pp.1-3, Dec 2005.

[11] R. A. Johnson and E. L. Titlebaum, "Range Doppler Uncoupling in the Doppler Tolerant Bat Signal", *Proc. of IEEE Ultrasonics Symposium* , New York, pp.64-67, 1972.

EXPERIMENTAL PATH LOSS MODELS FOR WIRELESS SENSOR NETWORKS

Rahul P. Sawant[†], Qilian Liang[†], Dan O. Popa[†*] and Frank L. Lewis[†*]
[†]Department of Electrical Engineering; [*]Automation & Robotics Research Institute
University of Texas at Arlington
Arlington, TX 76019 – 0016 U.S.A.
E – Mail: liang@uta.edu

ABSTRACT

Energy conservation is critical in Wireless Sensor Networks. Replacing or recharging batteries is not an option for sensors deployed in hostile environments. Generally communication electronics in the sensor utilizes most energy. This paper studies the effect of changing the transmission power and baud rate on transmission distance. Using Shannon channel capacity formula and Log – Distance Path Loss Model, transmission distance is shown to be related to transmit power and baud rate. Extensive empirical readings are taken to confirm the above relation. The path loss exponent got as a result of data fitting is within the acceptable range for wireless environment. Using the equation derived in this paper, the distance between neighboring motes and traffic density it will be possible for sensors to adjust their transmit power and baud rate so as to use only the required amount of energy to maintain the wireless link to the neighbor and conserve power.

I. INTRODUCTION

Wireless Sensor Networks (WSN) are comprised of small, inexpensive sensors with wireless communication capabilities, called motes. They are deployed in ad – hoc networks and are powered by limited power supplies. These motes are deployed in large numbers and provide unprecedented opportunities for instrumenting and controlling homes, cities and the environment. They find applications in different fields like military sensing, physical security, air traffic control, traffic monitoring, video surveillance, industrial automation etc. Each poses different challenges for these motes but one common challenge faced in all fields is power conservation. This is because motes are sometimes deployed in difficult to reach regions and this makes it difficult to replace the batteries. Hence power conservation becomes an important factor for these motes. One of the main reasons for deploying these motes in ad – hoc is power conservation. Power is consumed during data processing and RF communication, but communication electronics uses far more power than processing. Hence a lot of effort goes into designing energy efficient routing algorithms [1, 2, 3], directed diffusion algorithms [4, 5], clustering algorithms [6, 7, 8], data aggregation [9] and MAC protocols [10, 11, 12]. Many of these algorithms assume fixed transmit power and baud rate.

In this paper we study the effect of changing the transmit power and baud rate on transmission distance. This study is particularly useful for WSN's which are mobile, such as in applications involving robotic swarms [19]. We experimentally determine a relation between distance, transmit power and baud rate for Crossbow's Mica2 motes deployed in indoor environment. This can be incorporated along with the above algorithms and with knowledge of the distance to neighboring mote and the traffic density can be used to adjust the transmit power and baud rate so that packets can be forwarded to the neighboring mote with the least energy and maximum reliability. Adjusting the transmit power the neighboring mote will fall within the transmission range and neither will be out of range nor will be the transmission range much bigger than the desired distance. This will also help to keep packet exchange to the desired mote and not to any distant mote, which can itself start some other packet exchange and so will help keep interference to a minimum. This study along with the energy efficient algorithms will further increase energy efficiency in real time when deployed on Mica2 motes.

This paper is organized as follows. In section II we briefly describe the factors affecting transmission distance. Section III gives the details of our experiment and the empirical data plots obtained. In section IV curve and surface fitting is described and lastly in section V we provide a conclusion of our study.

II. TRANSMISSION DISTANCE

The study of Transmission Distance, Transmit Power and Baud Rate done for this paper is related to Energy Conservation in Wireless Sensor Network. Transmit Power and Baud Rate together affect the Transmission Distance. In a wireless channel there is some power loss called Path Loss. This path loss is inversely proportional to the distance between the transmitter and receiver. With the knowledge of distance and path loss we can find the exact transmit power required to maintain a good link to the receiver without wasting any energy.

1-4244-1513-06/07/$25.00

Another factor that affects the transmission distance is the antenna sensitivity. Baud rate of the data affects the receiver antenna sensitivity. This is because at the higher baud rate, there is less energy and a fewer number of actual radio waves in each bit of information transmitted. The lower baud rates have more energy per bit and more actual radio waves per bit transmitted. The more radio waves received per bit makes it much easier to establish the correct waveform. Thus, lower baud rates mean more energy per bit; better receive signals and longer transmission distances [13]. The advantage is by keeping the transmit power constant and only decreasing the baud rate we can increase the transmission distance of the radio.

III. Experiment Details

The following hardware was used for the experiment

1. MIB510 Programming board manufactured by Crossbow Technology.
2. Mica2 motes 2 nos. manufactured by Crossbow Technology
   For the radio communications Mica2 uses the CC1000 Chip which works in the 433 MHz band. The usual antenna chosen is a length of insulated wire called the monopole whip antenna one – quarter wavelength long. This happens to be 6.8 inches for the 433 MHz Mica2 [14]. Mica2 is powered by two AA batteries and requires a voltage between 2.7 – 3.3 Volts for successful operation.

A. CC1000 transceiver details

CC1000 uses the Binary Frequency Shift Keying (BFSK) modulation in the physical layer. Its transmit power and the baud rate are completely programmable. The crystal oscillator connected has a frequency of 14.7456 MHz. There are 22 inbuilt 8 – bit registers which can be used to program the operating parameters. A few relevant parameters are as follows.

Frequency Control Word A which sets the local oscillator frequency in receive mode and Frequency Control Word B which sets the transmitting frequency, f_0 in the transmit mode, together operate the frequency synthesizer (PLL). In the experiment Word A had a value of 0x580000 and B had a value of 0x57F685. Two registers FSEP0 and 1 are used to set the frequency separation and they hold a value of 0x355. Transmit and receive frequencies can be calculated using the following formula [15].

$$f_{VCO} = f_{ref} \times \frac{FREQ + (FSEP \times TXDATA) + 8192}{16384} \quad (1)$$

Here f_{VCO} gives the Local Oscillator (LO) frequency in receive mode and the f_0 and f_1 frequency in transmit mode (lower and upper FSK frequency). f_{ref} is the reference frequency calculated using the formula given below. *FREQ* is the value in frequency control word A or B according to whether the formula is used to calculate receive or transmit frequency. *FSEP* is the value in the FSEP0 and 1 registers. *TXDATA* is 0 or 1 in transmit mode depending on the data bit to be transmitted. In the receive mode *TXDATA* is always 0. Now to calculate f_{ref} we use the following formula [15]

$$f_{ref} = \frac{f_{xosc}}{REFDIV} \quad (2)$$

f_{xosc} is the crystal oscillator frequency of 14.7456 MHz and *REFDIV* is set to 12. This gives f_{ref} = 1.2288 MHz.

Using all these values in eq. 1, in receive mode $f_{VCO} = 433.152 \times 10^6 Hz$ and in transmit mode $f_{VCO} = 432.96997 \times 10^6 Hz$.

Thus the LO frequency in the receive mode is 433.152 MHz. In the experiments we have used a high side LO injection. So $f_{VCO} = f_{RF} + f_{IF}$ where f_{RF} is the centre frequency and f_{IF} is the Intermediate frequency. For CC1000 f_{IF} is designed to be 150.0375 KHz. So we get f_{RF} = 433.0019625 MHz.

The upper FSK transmit frequency is given by $f_1 = f_0 + f_{sep}$. f_0 is the f_{VCO} frequency calculated above for transmit mode i.e. 432.96997 MHz and f_{sep} is the frequency separation calculated using the formula [15] given below.

$$f_{sep} = f_{ref} \times \frac{FSEP}{16384} \quad (3)$$

Substituting the values $f_{sep} = 63975 Hz$, further using this f_{sep} we get f_1 = 433.03395 MHz. From f_0 and f_1 the centre frequency will be 433.0019625 MHz and is same as in the receive mode.

The transmit power of CC1000 can be varied from -20 dBm to 10 dBm in steps of 1 dB and is controlled by the PA_POW register. We carried out our experiments at four different power level setting as shown in Table 1.

Another parameter that was required for our experiments is the baud rate. This is set using a part of MODEM0 register as shown in Table 2. For our experiments we used 19.2, 38.4 and 76.8 kBaud.

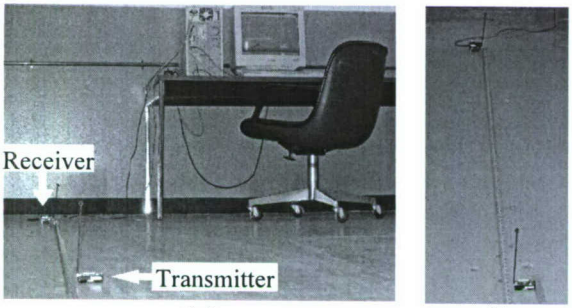
Table 1: PA_POW register values and corresponding output power [15]

Output Power (dBm)	PA_POW Register Value (hex)
-20	01
-19	01
-18	02
-17	02
-16	02
-15	03
-14	03
-13	03
-12	04
-11	04

Table 2: MODEM0 register part to set Baud Rate[15]

Register Part Name	Bit & Parameter Value
MODEM0{6:4} BAUDRATE{2:0}	000: 0.6 kBaud 001: 1.2 kBaud 010: 2.4 kBaud 011: 4.8 kBaud 100: 9.6 kBaud 101: 19.2, 38.4 and 76.8 kBaud 110: Not Used 111: Not Used

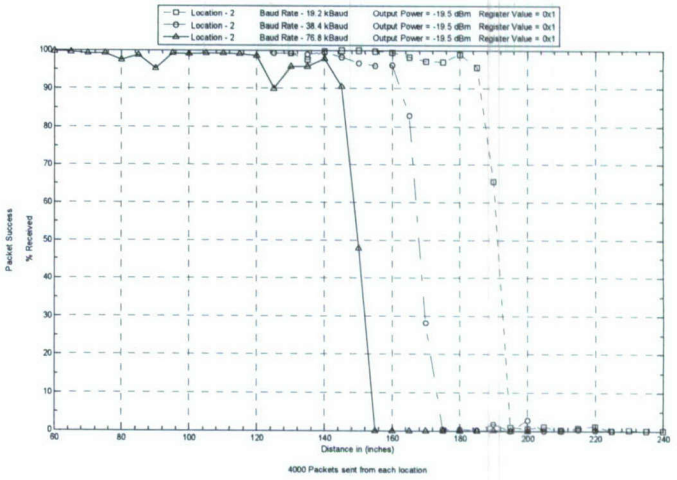
The experiment was carried out in a big warehouse which was used as a lab with minimal office furniture. The Mica2 motes were placed on the ground and were always in line of sight of each other. The following picture was taken during the experiment.



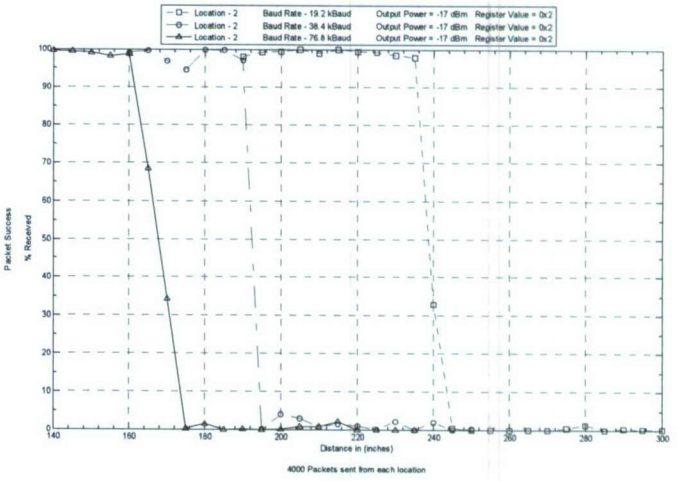
(a) (b)

Figure 1: Experimental Site (a) Front view (b) Top view of motes

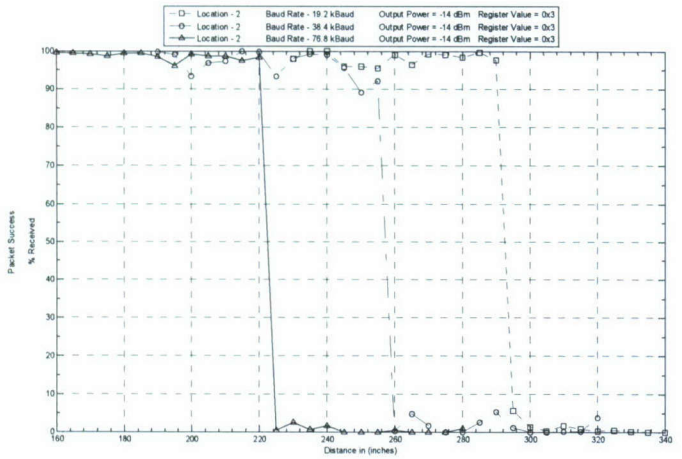
For the transmitter mote a certain power level and baud rate is set from the possible options available. Since the wireless channel varies a lot randomly with respect to time there will be some packet loss even when the receiver mote is within the transmission distance. Hence the transmitter mote sends a total of 4000 packets, so as to follow Monte Carlo method, each with a sequence number. The receiver mote keeps a record of the number of packets received and missed and finally transmits a summary packet to the computer attached to it. A resolution of 5 inches was used for the distance between the transmitter and receiver. Once it is found that the receiver is not receiving any packets at all we know that the receiver is out of the transmission range. When it is absolutely sure the receiver is out of range the transmit power or baud rate on the transmitter mote are changed to a different value and the above experiment carried again. The plots in Figure 2 were generated using the data gathered from the experiment.



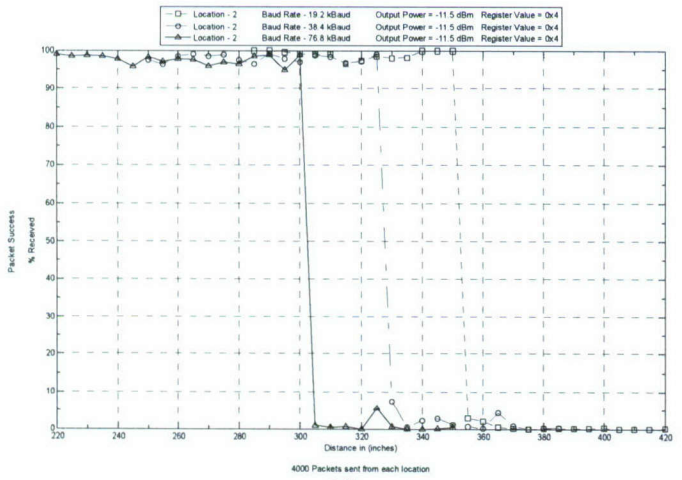
(a)



(b)



(c)



(d)

Figure 2: Packet success in percentage as a function of transmission distance for power levels of (a) -19.5 dBm, (b) -17 dBm, (c) -14 dBm, (d) -11.5 dBm and for 19.2 kBaud, 38.4 kBaud and 76.8 kBaud

Each curve from the plots in Figure 2 is initially in the region of 90 – 100 % this is because the transmitter and receiver motes are within the transmission range. A few packets are dropped but this could be because of variations in the channel that occur for a very small amount of time and is normal for any wireless channel. As the distance between the transmitter and receiver is increased, after a certain distance the curve drops drastically and the packets received are in the range of 0 – 5 %. This implies the transmitter is out of range of the receiver mote.

IV. Data Fitting

The empirical data is thought to be derivable from some underlying function. The plots in section III seem to closely follow the Q – function. We used curve fitting using method of least squares to minimize the sum of squared deviations. After fitting Q – function to each curve following plots were generated. All curve fits are not shown in Figure 3 but the data got from all curve fits is shown in Table 3.

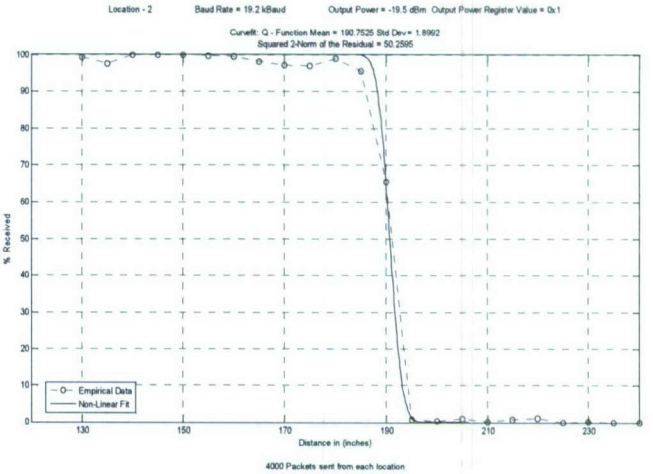


Figure 3: Q–function curve fit at 19.2 kBaud & -19.5 dBm

An initial value of one was used for std. dev. for curve fit and for mean of distance a value where packet success equals 50% was used. The following table shows the parameters got from the curve fit in above plots.

Table 3: Mean, Std. Dev. and Residual Error of curves fit above for different baud rate and transmit power

Baud Rate (kBaud)	Transmit Power (dBm)	Curve Parameters (Inches)		Residual Error
76.8	-19.5	Mean =	149.6819	208.8363
		Std Dev =	3.093	
	-17.0	Mean =	167.642	78.0394
		Std Dev =	4.567	
	- 14.0	Mean =	222.2506	41.3694
		Std Dev =	1.0526	
	- 11.5	Mean =	302.4448	148.6817
		Std Dev =	1.0754	
38.4	-19.5	Mean =	168.0964	57.4137
		Std Dev =	3.2833	
	-17.0	Mean =	191.759	79.4251
		Std Dev =	0.9316	
	- 14.0	Mean =	256.6509	324.2136
		Std Dev =	1.1637	
	- 11.5	Mean =	328.0941	115.0494
		Std Dev =	1.3191	
19.2	-19.5	Mean =	190.7525	50.2595
		Std Dev =	1.8992	
	-17.0	Mean =	239.1081	11.9331
		Std Dev =	2.0253	
	- 14.0	Mean =	292.7798	80.6525
		Std Dev =	1.3931	
	- 11.5	Mean =	353.1259	35.7488
		Std Dev =	0.9846	

For all the above std. dev. of distance the average = 1.899 in. The maximum error is 2.668 in. This is the error in the std. dev. of Q – function and not in the transmission distance. In transmission distance the error will be $2.668/2 = 1.334$ in. This is because only the distance between mean and monotonically decreasing part of Q – function is included while measuring transmission distance.

The mean of the distance in Table 3 is related to channel bandwidth and frequency of operation apart from transmit power and baud rate. This is three dimensional and we use surface fit as mentioned below.

A very important result derived by Shannon [16] is

$$C = B_c \times \log_2\left[1 + \frac{P_r}{(N_0 \times B_c)}\right] \quad (4)$$

Here C is the channel capacity in bits per seconds, B_c is the channel bandwidth in Hz, P_r is the received power in watts and N_0 is the single – sided noise power density in watts / Hz. Writing in terms of P_r and converting it to dBm we get

$$P_r(dBm) = 10 \times \log_{10}\left[1000 \times N_0 \times B_c\left(2^{\frac{C}{B_c}} - 1\right)\right] \quad (5)$$

In the wireless channel the received power is calculated using the formula [17] below.

$$P_r(d)[dBm] = P_t[dBm] - PL(d)[dB] \quad (6)$$

Here $P_r(d)$ is the received power in dBm at a distance d from transmitter, P_t is the transmit power in dBm and $PL(d)$ is the path loss in dB at a distance d from transmitter. For finding the path loss $PL(d)$ we use the Log – Distance Path Loss model [17] where

$$PL[dB] = PL(d_0) + \left\{10 \times n \times \log_{10}\left(\frac{d}{d_0}\right)\right\} \quad (7)$$

In this formula $PL[dB]$ is the path loss in dB, $PL(d_0)$ is the path loss at a reference distance d_0, n is the path loss exponent, d is distance in meters at which the path loss is to be calculated and d_0 is the reference distance also in meters.

Substituting eq. (7) in (6) and the result in (5) we get

$$P_t[dBm] - PL(d_0) - \left\{10 \times n \times \log_{10}\left(\frac{d}{d_0}\right)\right\} = 10 \times \log_{10}\left[1000 \times N_0 \times B_c\left(2^{\frac{C}{B_c}} - 1\right)\right]$$

Solving for d we have

$$d = d_0 \times 10^{\frac{P_t - PL(d_0) - 10 \times \log_{10}\left[1000 \times N_0 \times B_c\left(2^{\frac{C}{B_c}} - 1\right)\right]}{10 \times n}} \quad (8)$$

Thus we relate the transmission distance with transmit power, baud rate and channel bandwidth. We use this equation in the surface fit using method of least squares.

In the above equation $PL(d_0)$ is calculated using the Friis free space path loss formula [17] given below.

$$PL(d_0) = -10 \times \log_{10}\left[\frac{\lambda^2}{(4 \times \pi)^2 \times d_0^{\,2}}\right] \quad (9)$$

Here λ is the wavelength and d_0 is the reference distance. Centre frequency of the motes is 433.002 MHz, as calculated in section III, from which we get λ and the distance d_0 it is said should be in the far – field region of the transmitting antenna. This far – field region or Fraunhofer region of a transmitting antenna is defined as the region beyond the far – field distance d_f which is related to the antenna dimension and wavelength. It is given as follows [17]

$$d_f = \frac{2 \times D^2}{\lambda} \quad (10)$$

Where d_f is the far – field distance, D is the largest physical linear dimension of the antenna and λ is the wavelength.

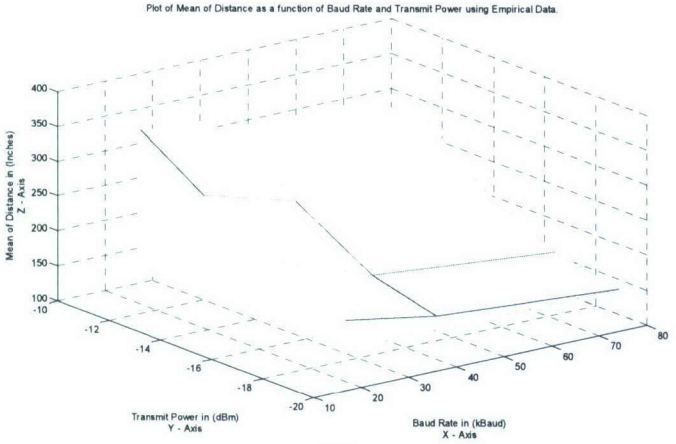
The Mica2 motes have a monopole whip antenna which is $\lambda/4$ meters high. Using 433.002 MHz frequency we get $\lambda = 0.6928$ m = 27.2767 in and the whip antenna height = 0.1732 m = 6.8 in. So D = 0.1732 m. Putting these values in eq. (10) we get $d_f = 0.086116$ m = 3.3903 in

Additional criteria to be in the far – field region are d_f must satisfy

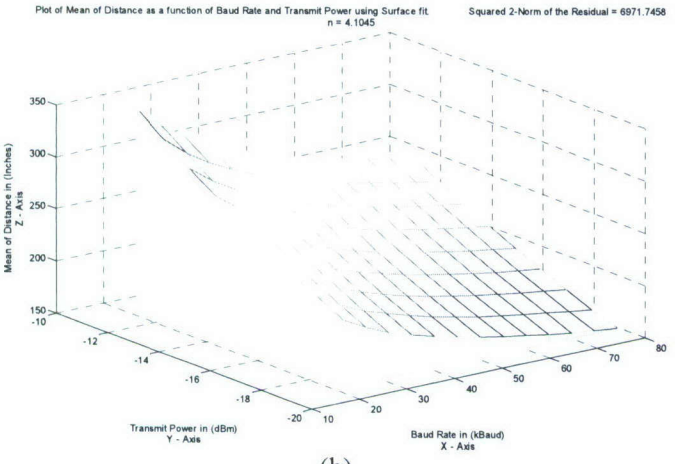
$d_f \gg D$ i.e. $d_f \gg 6.8$ in and

$d_f \gg \lambda$ i.e. $d_f \gg 27.2767$ in.

So to be in the far – field region we choose our distance d_0 to be 100 in = 2.54 m and from the λ value got above we calculate $PL(d_0)$ and use it in eq. (8). The N_0 in eq. (8) is calculated using $N_0 = k \times T$. k is the Boltzmann constant = $1.3806503 \times 10^{-23}$ m^2 kg s^{-2} K^{-1} and T is the room temperature in degree kelvin = 300^0 k. The channel bandwidth for 433 MHz Mica2 mote, $B_c = 620kHz$ [18]. Using all these values in eq. (8) we try to find the best values for n the path loss exponent such that the error, between the distances calculated using eq. (8) and the values of distance we got from our experimental data, is minimized. To start with, a value of one is assumed for n. After doing the surface fit we get the following graphs.



(a)



(b)

Figure 4: Mean of Distance vs Transmit Power and Baud Rate using (a) Empirical Data (b) eq. (8) derived from Shannon theorem and Path Loss Model

The optimized value of path loss exponent n got from surface fit is n = 4.10. The following graph compares distances using both data and also shows the surface.

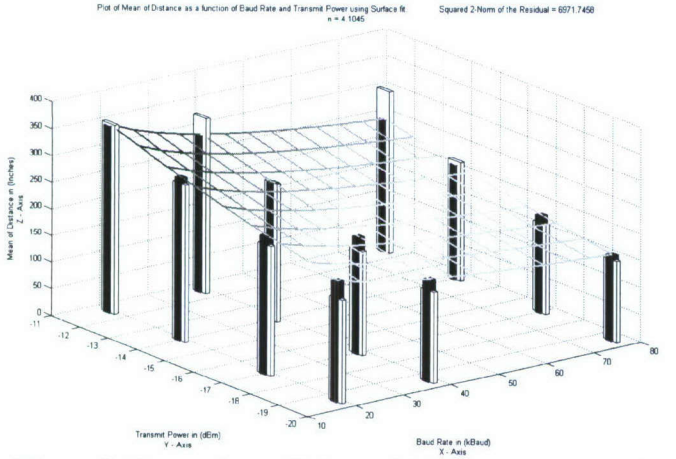


Figure 5: Comparison of Mean of Distance vs. Transmit Power & Baud Rate using Empirical & Surface Fit data

In the plot above the white bars show the distances got using the empirical data at baud rates of 19.2, 38.4 and 76.8 kBaud. The black bars show the distance got using eq. (8). So the difference between the two is the error. For comparison purpose the table below shows the empirical distance, distance obtained using eq. (8), the absolute error between the two and the percentage error.

Table 4: Summary of mean of distance and absolute error

Empirical Data			
Transmit Power	Baud Rate		
	19.2 kBaud	38.4 kBaud	78.6 kBaud
- 19.5 dBm	190.7526	168.0964	149.6819
-17.0 dBm	239.1081	191.7589	167.6420
-14.0 dBm	292.7798	256.6509	222.2506
-11.5 dBm	353.1259	328.0941	302.4448
Shannon Formula Surface Fit - n = 4.10			
Transmit Power	Baud Rate		
	19.2 kBaud	38.4 kBaud	78.6 kBaud
- 19.5 dBm	223.0249	187.8755	157.8458
-17.0 dBm	256.6032	216.1618	181.6109
-14.0 dBm	303.6356	255.7817	214.8980
-11.5 dBm	349.3505	294.2918	247.2528
Absolute Error			
Transmit Power	Baud Rate		
	19.2 kBaud	38.4 kBaud	78.6 kBaud
- 19.5 dBm	32.2723	19.7791	8.1639
-17.0 dBm	17.4952	24.4028	13.9689
-14.0 dBm	10.8558	0.8692	7.3526
-11.5 dBm	3.7754	33.8022	55.1921
Max Error - 55.1921		Min Error - 0.8692	
Percentage Error			
Transmit Power	Baud Rate		
	19.2 kBaud	38.4 kBaud	78.6 kBaud
- 19.5 dBm	16.9184	11.7665	5.4542
-17.0 dBm	7.3168	12.7258	8.3326
-14.0 dBm	3.7078	0.3387	3.3083
-11.5 dBm	1.0691	10.3026	18.2486

Note: All distances in Inches.

The error in the above table is very small. A close comparison of empirical data and surface fit data reveals a mix of positive and negative errors in mean of distance which will reflect in the same way in transmission distance. When this equation is used in a networking protocol most of the time the transmit power and baud rate set will maintain a good link to the receiver. The bad link occurring, in spite of setting the transmit power and baud rate as per eq(8), can happen also because of the dramatic change in the environment for short or long period of time. In this case the transmit power could be slightly increased or the baud rate slightly decreased to maintain a good link. Once done these new values can be used throughout the lifetime of the network.

Tuning of the transmission power and baud rate will only increase the network setting time a little but afterwards throughout the operation of the network energy will be conserved. The equation in this paper definitely gives a better starting point at fixing the transmit power and baud rate rather than approximately varying them from minimum to maximum and wasting a lot of time for network setup.

V. CONCLUSION

The key challenge in Wireless Sensor Networks being power conservation, the study in this thesis will support multi – hop routing and clustering algorithms in increasing the power saving in each sensor node. Using Shannon theorem and Log – Distance Path Loss Model the transmission distance is shown to be related to transmit power and baud rate. The extensive empirical data obtained confirms to this relation of transmission distance. The optimized value of the path loss exponent obtained as a result of surface fitting is close to that obtained in other wireless environments. Using this relation along with the knowledge of localization and traffic density between the neighboring mote it will be possible to adjust the transmit power and / or the baud rate to maintain the wireless links between the neighboring motes without loosing the connection and without incurring huge energy costs.

Apart from the advantage of power saving, the correct transmit power level will decrease the level of interference in the network since less motes in the surrounding will hear the conversation. There is also an increase in the network capacity as the packet transmissions are confined only to the small local area and other motes out of the area are free to carry out their own transmissions with some other motes within their area.

ACKNOWLEDGMENT

The authors would like to thank Prasanna M. Ballal of ARRI's DIAL Lab for his help with the experimental testbed used in this paper.

This work was supported in part by Office of Naval Research (ONR) under Grant N00014 – 07 – 1 – 0395, N00014 – 07 – 1 – 1024, and National Science Foundation (NSF) under Grant CNS – 0721515.

REFERENCES

[1] A. Woo, T. Tong and D. Culler, "*Taming the Underlining Challenges of Reliable Multihop Routing in Sensor Networks*" Proc. of 1st Intl. conf. on Embedded Networked Sensor Systems SenSys November 2003.

[2] C. Perkins and E. Royer, "*Ad hoc On – Demand Distance – Vector Routing*". Proc. of 2nd IEEE Workshop on Mobile Computing Systems and Applications, 1999.

[3] D. Braginsky, D. Estrin, "*Rumor Routing Algorithm For Sensor Networks*" Proc. of 1st Intl. Workshop on WSN & App., Sept 2002, Pages: 22 – 31.

[4] C. Intanagonwiwat, R Govindan, D Estrin, J Heidemann, F Silva, "*Directed Diffusion for Wireless Sensor Networking*" IEEE/ACM Trans. on Networking, Vol. 11, No. 1, February 2003 Pages: 2 – 16.

[5] J. Kulik, W. Rabiner and H. Balakrishnan, "*Adaptive Protocols for Information Dissemination in Wireless Sensor Networks*" Proc. 5th Annu. ACM/IEEE Int. Conf. Mobile Computing and Networking (MobiCom 1999), Pages: 174 – 185.

[6] W. Heinzelman, A. Chandrakasan and H. Balakrishnan, "*Energy – Efficient Routing Protocols for Wireless Microsensor Networks*" Proc. of 33rd Hawaii Int. Conf. System Sciences (HICSS), Maui, HI, January 2000.

[7] W. Heinzelman, A. Chandrakasan and H. Balakrishnan, "*An Application – Specific Protocol Architecture for Wireless Microsensor Networks*" IEEE Trans. Wireless Comm., Vol. 1, No. 4, October 2002 Pages: 660 – 670.

[8] A. Wang, W. Heinzelman, and A. Chandrakasan, "*Energy – Scalable Protocols for Battery – Operated Microsensor Networks*" Proc. 1999 IEEE Workshop Signal Processing Systems (SiPS 1999), October 1999, Pages: 483–492.

[9] B. Krishnamachari, D. Estrin, S. Wicker, "*The Impact of Data Aggregation in Wireless Sensor Networks*", Proc. of 22nd Intl. Conf. on Distributed Computing Systems, July 2002, Pages: 575 – 578.

[10] J. Polastre, J. Hill and D. Culler, "*Versatile Low Power Media Access for Wireless Sensor Networks*" Proc. of 2nd Intl. Conf. on Embedded Networked Sensor Systems, November 2004, Pages: 95 – 107.

[11] W. Ye, J. Heidemann, D Estrin, "*An Energy Efficient MAC protocol for Wireless Sensor Networks*" Proc. IEEE Infocom 2002, Vol. 3, Pages: 1567 – 1576.

[12] Q. Ren, Q. Liang, "*A Contention Based Energy – Efficient MAC Protocol for Wireless Sensor Networks*", IEEE Wireless Communications and Networking Conf., Vol. 2, April 2006, Pages: 1154 – 1159.

[13] "*Wireless Instrumentation: Factors Affecting Transmission Distance*", Accutech Tech Note # 215, Accutech Instrumentation Solutions, http://www.savewithaccutech.com/tech-center/technical-articles.asp

[14] "*MPR / MIB User's Manual*", Rev. A September 2005, Crossbow Technology, http://www.xbow.com/Support/wUserManuals.aspx.

[15] "CC1000 Single Chip Very Low Power RF Transceiver Datasheet", Chipcon Corporation, http://www.chipcon.com/files/CC1000_Data_Sheet_2_3.pdf

[16] J. Proakis, "*Digital Communication*", Fourth Edition, McGraw – Hill, 2001.

[17] T. Rappaport, "*Wireless Communications: Principles and Practice*", Second Edition, Pearson Education, 2004.

[18] J. Jeong, S. Kim, "*DOT3 Radio Stack*", www.cs.berkeley.edu/~jaein/presentations/JaeinJeong_ChipconRadioStack.ppt

[19] D.O. Popa and F.L. Lewis, "*Algorithms for robotic deployment of WSN in adaptive sampling applications*" in Wireless Sensor Networks and Applications, ed. Y. Li, M. Thai, and W. Wu, Springer-Verlag, Berlin, 2006, to appear.

Distributed Connected Dominating Set Construction in Random Geometric k-Disk Graphs

Dechang Chen
Uniformed Services University
Of the Health Sciences
Bethesda, MD 20814, USA
dchen@usuhs.mil

Kai Xing
Computer Science
The George Washington University
Washington, DC 20052, USA
kaix,wcheng@gwu.edu

Qilian Liang
Electrical Engineering
University of Texas at Arlington
Arlington, TX 76019, USA
liang@uta.edu

E. K. Park
Computer Science & Electrical Engineering
University of Missouri at Kansas City
Kansas City, MO 64110, USA
ekpark@umkc.edu

***Abstract*—In this paper, we study the problem of minimum connected dominating set in random geometric k-disk graphs. This research is motivated by the problem of virtual backbone construction in wireless ad hoc and sensor networks, where the coverage area of nodes are disks with different radii. We derive the size relationship of any maximal independent set and minimum connected dominating set in geometric k-disk graphs, and apply it to analyze the performances of two distributed connected dominating set algorithms we propose in this paper. These algorithms have bounded performance ratio and low communication overhead, and therefore have the potential to be applied in real ad hoc and sensor networks.**

***Index Terms*—connected dominating set, random geometric k-disk graphs, maximal independent set, performance ratio.**

I. Introduction

Let V be the set of points denoting the set of nodes randomly placed in the Euclidean plane. Associated with $\forall u \in V$ is a radius r_u such that $1 \leq r_u \leq k$, where k is a constant. Let $G_k(V, E)$ be the disk graph constructed from V such that an edge $(u, v) \in E$ if and only if the distance between u and v is at most $\min\{r_u, r_v\}$. We call G_k a *random geometric k-disk graph* or a *geometric k-disk graph*. When $k = 1$, G_k is a unit-disk graph.

Geometric k-disk graphs have been widely adopted to model wireless ad hoc and sensor networks, in which two nodes can communicate with each other successfully via two-way handshake (DATA-ACK) if they reside in each other's transmission range. Since their power capacity differs or for the reason of interference mitigation, nodes may have different transmission ranges, resulting in disks with different radii.

In this paper we study the problem of constructing minimum connected dominating sets (MCDS) in geometric k-disk graphs. One significant application of this problem is the virtual backbone construction in ad hoc and sensor networks, where a virtue backbone is utilized to decrease the protocol overhead [1], [2]. Computing a MCDS in $G_k(V, E)$ is a NP-hard problem since its special case, MCDS in unit-disk graphs, is NP-hard [3]. In this paper, we first figure out the size relationship between any maximal independent set (MIS) and the minimum connected dominating set in G_k. Then we report two distributed approximation algorithms for connected dominating set construction.

Our major contributions are three folds. First, we prove that the maximum number of independent neighbors a node may have in a geometric k-disk graph is at most $5 + 9\lceil \frac{\ln k}{\ln(2\cos(\frac{\pi}{5}))} \rceil$. Second, we derive the size relationship between any MIS and the MCDS in geometric k-disk graphs, which is $(7\lceil \frac{\ln k}{\ln(2\cos(\frac{\pi}{5}))} \rceil + 4)opt + 2\lceil \frac{\ln k}{\ln(2\cos(\frac{\pi}{5}))} \rceil + 1$, where opt is the size of any MCDS. Note that the best results in literature were reported by Thai et al. in [4], which stated that when $k > 1$, each node has at most $10\lfloor \frac{\ln k}{\ln(2\cos\frac{\pi}{5})} \rfloor$ number of independent neighbors and the size of any MIS is at most $10\lfloor \frac{\ln k}{\ln(2\cos\frac{\pi}{5})} \rfloor opt$. Notice that our results work for unit-disk graphs too, where $k = 1$.

The third contribution of this paper is the two distributed approximation algorithms, for which we conduct rigorous theoretical performance analysis. In the first algorithm, we grow a CDS from a leader. This algorithm requires each node know the ids of its neighbors. In the second algorithm, we construct an MIS first then find out connectors to connect all nodes in the MIS. This algorithm takes node degree into consideration and therefore has higher message overhead but achieves better performance compared to the first one.

This paper is organized as follows. We first introduce the preliminary definitions in Section II. Related works are also surveyed in this section. Then we derive the size relationship between maximal independent set and minimum connected dominating set in geometric k-disk graphs in Section III. The two distributed approximation algorithms are reported in sections IV and V, respectively. We conclude this paper with a brief summary and a brief discussion on future research in Section VI.

II. Preliminaries And Related Work

A. Preliminaries

Given any graph $G = (V, E)$, two vertices are *independent* if they are not neighbors. For $\forall\ u, v \in V$, $hop\_count(u, v)$ is the number of edges (hops) in the shortest path from u to v. For any vertex v, $N_1[v] = \{u \mid hop\_count(u, v) \leq 1\}$ is the *one-hop (close) neighbor set* of v; $N_2[v] = \{u \mid hop\_count(u, v) \leq 2\}$ is the *two-hop (close) neighbor set* of v.

An *independent neighbor set* of v, denoted by $N_I(v)$, is a subset of $N_1[v]$ such that any pair of vertices in $N_I(v)$ are independent. An *independent set* S of G is a subset of V such that for $\forall\ u, v \in S$, $(u, v) \notin E$. S is *maximal* if any vertex not in S has a neighbor in S.

A *dominating set* D of G is a subset of V such that each node not in D has at least one neighbor in D. For any edge (u, v), if $u \in D$ and $v \notin D$, then u is v's *dominator* and v is u's *dominatee*. If both u and v are $\in D$, one can specify the other as its dominator. An optimal dominating set has minimum cardinality. If the induced subgraph of D is connected, then D is a *connected dominating set (CDS)*. Among all CDSs of graph G, the one with minimum cardinality is called a *minimum connected dominating set (MCDS)*. A maximal independent set is also a dominating set.

B. Related Work

To our best knowledge, Thai et al. [4] was the first and the only one to tackle the problem of MCDS construction in geometric k-disk graphs. In that work, the authors proved that that any node can have at most $10\lfloor\frac{\ln k}{\ln(2\cos\frac{\pi}{5})}\rfloor$ number of independent neighbors when $k > 1$. This result is problematic as when k is a little more than 1, it yields 0, which is impossible. Thai et al. [4] also proposed three centralized approximation algorithms together with their performance analysis. Our work is motivated by [4] but improves [4]. In addition, we focus on the design of distributed algorithms for the geometric k-disk graphs.

There exist abundant works investigating the problem of MCDS construction in unit-disk graphs and general graphs. For a recent literature survey, we refer the readers to [1] and the references therein. In the following we briefly summarize several major works.

The NP-Completeness of MCDS in general graphs was studied in [5]. The MCDS remains NP-hard for unit-disk graphs [3]. In 1998, Guha and Khuller proposed two CDS construction strategies in their seminar work [6]. These two greedy heuristic algorithms have performance ratios $2(1 + H(\Delta))$, where Δ the largest node degree and H is the harmonic function, and $3 + \ln\Delta$, respectively. Ruan *et al.* [7] proposed a one-step greedy approximation algorithm with a performance ratio of $2 + \ln\Delta$. Wu and Li [8] proposed the first localized algorithm for MCDS in general graphs but their algorithm does not have a bounded performance guarantee [9].

For unit-disk graphs, Wu *et al.* [10] studied the size relationship between MCDS and MIS and reported that the size of any MIS is at most 3.8 times of that of MCDS plus a constant 1.2. With this result, the performance ratio for the approximation algorithms in [9], [11], [12], which connect a MIS with a spanning tree, is improved from 8 to 7.8. Li *et al.* [13] designed the algorithm of connecting an MIS with a Steiner tree, and achieved a performance ratio of $5.8 + \ln 4$. Note that all the above-mentioned works require a MIS satisfying the follwing property: any subset of the MIS is two-hop away from its complementary. The construction of such a MIS relies on a spanning tree. The works by Wan *et al.* [14] and Cheng *et al.* [2] removed this restriction with a tradeoff of a larger performance ratio. MCDS in unit-disk graphs has a polynomial time approximation scheme [15], which means that MCDS in unit-disk graphs can be approximated to any degree.

The research reported in this paper is motivated by [2] and [4] and we target the geometric k-disk graphs. To our knowledge, the results reported in this paper are the best in literature.

III. Maximal Independent Set and Minimum Connected Dominating Set in Geometric k-disk Graphs

Given $G_k(V, E)$, a geometric k-disk graph as defined in Sec I, let $N_I(x)$ be the set of independent neighbors of x for for $\forall x \in V$. We have

Lemma 3.1: $|N_I(x)| \leq 5 + 9\lceil\frac{\ln k}{\ln(2\cos(\frac{\pi}{5}))}\rceil$.

Proof: Denote by r_x the radius of x. Let α be a real number such that $0 < \alpha < \frac{\pi}{3}$. For easier elaboration we assume that $\frac{2\pi}{\alpha}$ is a constant and we set $\alpha = \frac{\pi}{5}$. We will explain at the end of the proof that this α value is a reasonable choice. Note that the whole proof procedure only requires $0 < \alpha < \frac{\pi}{3}$.

We first draw circles centered at x with radii $(2\cos\alpha)^0$, $2\cos\alpha$, $(2\cos\alpha)^2$, $\cdots$, $(2\cos\alpha)^{n_s-1}$, and $(2\cos\alpha)^{n_s}$, respectively, such that $(2\cos\alpha)^{n_s-1} < r_x$ and $(2\cos\alpha)^{n_s} \geq r_x$. Therefore $n_s = \lceil\frac{\ln r_x}{\ln(2\cos\alpha)}\rceil$. By this way the disk of x is partitioned into n_s annuluses plus the unit-radius circle centered at x, as shown in Fig. 1. We claim that each annulus contains at most $\frac{2\pi}{\alpha} - 1$ nodes in $N_I(x)$.

Consider any annulus $\mathcal{H}$, as shown in Fig. 1(a). Let u and v be two independent neighbors of x in $\mathcal{H}$. We must have $\angle uxv > \alpha$. To prove this claim, we draw two lines xb and xd crossing the inner and outer sides of $\mathcal{H}$ at a and b, and c and d, respectively, such that u resides in xb and $\angle bxd = \alpha$. Since $\frac{|xb|}{|xc|} = \frac{|xd|}{|xa|} = 2\cos\alpha$, we have $|cb| = |ad| = |xa| = |xc|$. Therefore $|yx| > |yb|$ holds true for any point y in the closed area $abdc$ of $\mathcal{H}$, which means all nodes in the closed area of $abdc$ in $\mathcal{H}$ are neighbors of u. Thus v must be out of the area of $abdc$, and therefore $\angle uxv > \alpha$. Based on this argument, we conclude that there exist at most $\frac{2\pi}{\alpha} - 1$ number of independent neighbors of x in the annulus $\mathcal{H}$.

Since there are n_s number of annuluses, the total number of independent neighbors of x in all the annuluses is at most $n_s \cdot [\frac{2\pi}{\alpha} - 1] = \lceil\frac{\ln r_x}{\ln(2\cos\alpha)}\rceil \cdot [\frac{2\pi}{\alpha} - 1]$. Let $f(\alpha) = \lceil\frac{\ln r_x}{\ln(2\cos\alpha)}\rceil \cdot [\frac{2\pi}{\alpha} - 1]$. When $0 < \alpha < \frac{\pi}{3}$, $f(\alpha)$ yields a minimum when α is close to $\frac{\pi}{5}$.

Now consider the unit disk $\mathcal{D}$ centered at x, as shown in Fig. 1(b). Let u and v be two independent neighbors of x. Then we must have $\angle uxv > 60°$, since otherwise, $|uv| \leq$

$\max\{|ux|, |vx|\} \leq 1$, contradicting the independence of u and v. Therefore there exist at most 5 independent neighbors of x at $\mathcal{D}$.

Considering the unit-disk area and all the annulus areas, we have $|N_I(x)| \leq 5 + 9\lceil \frac{\ln r_x}{\ln(2\cos(\frac{\pi}{5}))} \rceil \leq 5 + 9\lceil \frac{\ln k}{\ln(2\cos(\frac{\pi}{5}))} \rceil$. ■

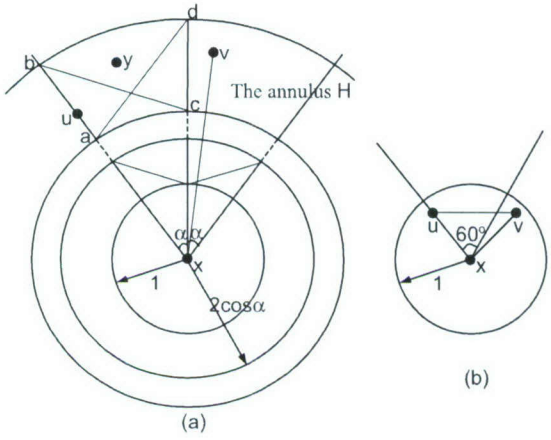


Fig. 1. (a) If u and v are independent neighbors in the annulus $\mathcal{H}$, then $\angle uxv > \alpha$. Note that only part of $\mathcal{H}$ is shown. (b) When u and v are independent neighbors in the unit-disk area of x, $\angle uxv > 60°$.

In the following analysis, we denote $\lceil \frac{\ln k}{\ln(2\cos(\frac{\pi}{5}))} \rceil$ by n_k. Now we have derived the upper bound of the number of independent neighbors for any node x in $G_k(V, E)$, which is $5 + 9n_k$. In the next, we are going to answer the following question: how many independent neighbors a pair of neighboring nodes may have? The following lemma answers this question by giving an upper bound.

Lemma 3.2: Let u and v be any pair of neighboring nodes in $G_k(V, E)$, i.e. $u \in V$, $v \in V$, and $(u, v) \in E$. Then the total number of independent neighbors of u and v is at most $9 + 16n_k$.

Proof: Let r_u and r_v be the radii of u and v, respectively. Without loss of generality, we assume $r_u \geq r_v$. Assume the disks of u and v intersect at x and y, as illustrated in Fig. 2. We have $|ux| = r_u$ and $|vx| = r_v$. Since $|ux| \geq |vx|$ and $|ux| \geq |uv|$, we have $\angle uvx \geq \frac{\pi}{3}$. Similarly we have $\angle uvy \geq \frac{\pi}{3}$. Therefore $\angle xvy \geq \frac{2\pi}{3}$. Thus the area in v's disk that could cover nodes in $N_I(v)$ but not in $N_I(u)$ is the right sector $\widehat{xvy}$ of v with an angle at most $\frac{4\pi}{3}$.

Now we can take the same procedure as that in Lemma 3.1 by drawing circles centered at v with radii $1, 2\cos\alpha, \cdots$. Since the angle at v formed by any two independent neighbors of v in the same annulus is larger than α, the overlapping area of the right sector $\widehat{xvy}$ and all the annuluses can contain at most $(\lceil \frac{\frac{4\pi}{3}}{\alpha} \rceil) n_k = 7n_k$ number of independent neighbors of v. With a similar argument, the intersection of the unit-disk centered at v and the right sector $\widehat{xvy}$ may contain at most $\frac{\frac{4\pi}{3}}{\frac{\pi}{3}} = 4$ independent neighbors. Therefore the total number of independent neighbors of u and v is at most $9 + 16n_k$. ■

Lemma 3.2 indicates that a neighbor can bring in at most $4 + 7n_k$ number of independent nodes. This motivates us to prove the following lemma addressing the size relationship between a minimum connecting dominating set and any maximal independent set of the same G_k.

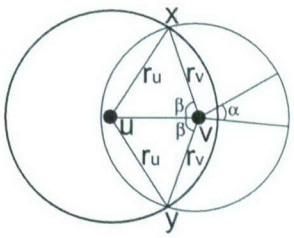


Fig. 2. If u and v are neighbors in $G_k(V, E)$ and $r_u \geq r_v$, then $\angle xvy > \frac{2\pi}{3}$.

Let $S \subset V$ be any MIS of $G_k(V, E)$. Denote by *opt* the size of any MCDS in G. We have

Theorem 3.1: $|S| \leq (7n_k + 4)opt + 2n_k + 1$.

Proof: Let C be any MCDS of $G_k(V, E)$. Compute a spanning tree τ traversing all nodes in C. Let $u_1, u_2, \cdots, u_{opt}$ be the preorder-tree-walk of τ. Then from Lemma 3.1 u_1 will dominate at most $5 + 9n_k$ number of independent nodes in S. Based on Lemma 3.2, each u_i, where $i = 2, 3, \cdots, opt$, will dominate at most $4 + 7n_k$ number of independent nodes in S that are not dominated by any node in u_j where $j < i$. Since C is a connected dominating set, each node in S will be either in C, or dominated by one node in C. Therefore $|S| \leq 5 + 9n_k + (opt - 1)(4 + 7n_k) = (4 + 7n_k)opt + 2n_k + 1$. ■

In the following sections, we propose two distributed CDS algorithms and analyze their performance theoretically. Note that for both algorithms, we assume that each node has a unique id, known to its neighbors within two hops away.

IV. Algorithm I: Growing a CDS from a Leader

Algorithm I starts from a leader node. For simplicity, we assume that the node with the smallest id is the leader. We associate a color with each node. Strictly speaking, *color* is not a parameter in our algorithm. It is retained in the algorithm description for the purpose of easier elaboration.

Initially all nodes are colored white. In the first step, leader u colors itself black and becomes a dominator. Then all nodes in $N_1[u] \setminus \{u\}$ color themselves gray and become dominatees, and all nodes in $N_2[u] \setminus N_1[u]$ color themselves yellow and become active. In the next step, neighboring active nodes compete with each other and the winners, whose ids are the smallest among their yellow neighbors, become dominators. Each of the winners also specifies its own dominator, the gray neighbor whose id is the smallest among all gray neighbors. This step will repeat until nodes are either gray (dominatees) or black (dominators). All black nodes form a CDS.

Note that the procedure described above grows a CDS from the leader. In the first step, only the leader is included in the CDS. In each of the other steps, two nodes are included in the CDS, with one turning color from yellow and one from grey.

Theorem 4.1: The size of the CDS generated by Algorithm I is at most $(14n_k + 8)opt + 4n_k + 1$, where *opt* is the size of the minimum connected dominating set.

Proof. We can partition all the dominators into two sets: A and B. Set A contains all vertices with color changing from *white to black directly* and B contains all vertices with color changing from *white to gray then to black*. The first step adds the leader to A. Each of the other steps adds one node to A and one node to B. Thus $|A| = |B| + 1$.

Now we claim that A is an independent set. This is obvious since each vertex u in A is colored black from white or yellow. This means u has no black neighbors because each neighbor of a black node has gray color. From Lemma 3.1, $|A| \leq (7n_k + 4)opt + 2n_k + 1$. Thus $|A| + |B| \leq (14n_k + 8)opt + 4n_k + 1$. ■

Remarks: (i) Algorithm I grows a tree from the leader in a step-by-step fashion. At any time, all the inner nodes of the tree are colored black while all the leaf nodes are colored gray. In the first step, the leader and all of its neighbors are added to the tree. In every other step, a leaf node v and one of its yellow neighbors u are colored black (added to the tree). All the white/yellow neighbors of u and v are colored gray and added to the tree as leaves. This algorithm terminates when no white/yellow node left. All the black nodes form the CDS. (ii) If we use node cost instead of id as the criteria for dominator selection, Algorithm I is cost-aware. For example, if the cost is the inverse of the residual power of each node, the output CDS has higher power capacity; if the cost is the incoming bitrate (load), the output CDS has lower load; if the cost is the inverse of the node velocity, the induced graph by the output CDS has a more stable topology.

A. Distributed Implementation

Each node u maintains the following parameters: dom_u, which is the dominator, or the parent of the node in the tree; $rank_u$, which defines a relative relationship among neighboring nodes[1]; $children_u$, which contains all dominated nodes, or the children of the node in the tree. These parameters are updated by the exchange of the following 3 messages.

$< dominator(u, dom_u, rank_u) >$ — node u, whose dominator is dom_u and whose rank is $rank_u$, broadcasts this message to all neighbors. u is a dominator.

$< dominatee(u, dom_u, rank_u) >$ — node u, whose dominator is dom_u and whose rank is $rank_u$, broadcasts this message to all neighbors. u is a dominatee.

$< active(u) >$ — node u broadcasts this message to all of its white/yellow neighbors when it becomes active. A white node becomes active after it receives the first $< dominatee >$ message from one of its neighbors.

The state transition diagram of Algorithm I is given in Figure 3. Each node u runs a copy of the algorithm. At any time, u can be in one of the 4 states: S_0, S_1, S_2, and S_3. The directed arc from S_i to S_j, where $i, j = 0, 1, 2, 3$, represents the transition from state S_i to state S_j. Each transition is labeled by a number. These transitions will be explained latter.

[1]In Algorithm I, rank is the level of the node in the tree

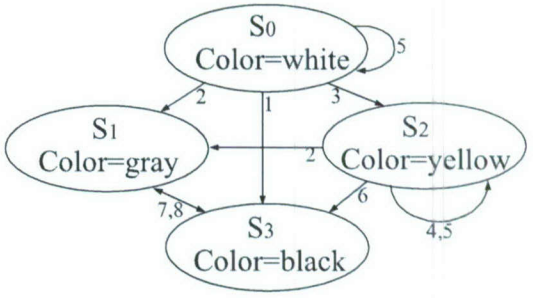


Fig. 3. *The state transition diagram of Algorithm I for any node u.*

State S_0 is the initial state. A node in this state has white color. All nodes are in S_0 at the beginning of the algorithm. State S_1 is the dominatee state. A node in this state is a dominatee and has gray color. State S_2 is the active state. A node in this state has at least one neighbor in S_1 and has yellow color. An active node is a candidate dominator in next step. State S_3 is the dominator state. A node in this state is a dominator and has black color. All nodes in S_3 form the connected dominating set. Each node u in state S_0 or S_2 also maintain a parameter W_u to node all yellow neighbors. Initially $W_u = \emptyset$. If u has the smallest id compared to its yellow neighbors in W_u, u will become a dominator in the next step. The transition steps are detailed below.

1) u is in state S_0. If u is the leader, then $dom_u = u$, $rank_u = 0$. u will broadcast message $< dominator(u, u, 0) >$ and go to state S_3.
2) u is in state S_0 or S_2 and receives message $< dominator(v, d, l) >$ from neighbor v. If d is a neighbor of u, then $dom_u = d$, $rank_u = l$; otherwise, $dom_u = v$, $rank_u = l + 1$. u will broadcast message $< dominatee(u, dom_u, rank_u) >$ and go to S_1.
3) u is in state S_0 and receives message $< dominatee(v, d, l) >$ from neighbor v. u will broadcast message $< active(u) >$ and go to state S_2. u temporarily sets dom_u to v and $rank_u$ to $l + 1$. If v is in W_u, it will be removed.
4) u is in state S_2 and receives message $< dominatee(v, d, l) >$. Remove v from W_u if it is in W_u. If $id_v < id_{dom_u}$, then $dom_u = v$ and $rank_u = l + 1$. u will go back to S_2.
5) u is in state S_0 or S_2 and receives message $< active(v) >$. $W_u = W_u \cup \{v\}$. u remains in the original state.
6) u is in state S_2 and there is no broadcasting in $N_1[u]$ for T_0 time unit (which is a design constant). If u has smallest id compared with all nodes in W_u, it will broadcast message $< dominator(u, dom_u, rank_u) >$ and go to S_3.
7) u is in state S_1 and receives message $< dominator(v, d, l) >$ from v. If $d = u$, u will broadcast message $< dominator(u, dom_u, rank_u) >$ and go to state S_3.
8) u is in state S_3. If u finds that none of its neighbors

takes u as the dominator, u will broadcast message $dominatee(u, dom_u, rank_u)$ and go to state S_1.

Remarks: (i) The leader can be elected by the distributed leader election algorithm proposed in [16]. (ii) The parameter T_0 is used to force the start of next step in nearby environment be strictly after the end of the current step. T_0 is a design parameter, which could be predetermined via simulation. (iii) We assume the message broadcast is reliable. For reliable broadcasting, we refer the readers to [17], [18]. (iv) When u receives a message $< dominator(v, u, l)$ or $dominatee(v, u, l)$, u will put v into its children list if v is not there. (v) In state S_1, a gray node will select the lowest rank black node as its dominator. This can help to optimize the generated CDS. For example, if a black node u's dominated children are also dominated by $dom(u)$, then u can become a dominatee and go to state S_1. (vi) Step 8 is an optimization to remove those dominators with no dominatees.

Theorem 4.2: Algorithm I has a message complexity $O(n)$, where n is the total number of vertices.

Proof. We have 3 types of messages: $< dominator >$, $< dominatee >$, and $< active >$. Each node broadcasts each message at most once. Thus the total number of broadcastings is at most $O(n)$. ■

Note that the above message complexity analysis does not consider the leader election, whose message complexity is $\Omega(n \log n)$ [14].

B. An example

Now we use an example to demonstrate how to apply Algorithm I to compute a connected dominating set. The given geometric disk graph $G = (V, E)$ is shown in Figure 4. There are 6 nodes and 11 links. Host 0 is the leader.

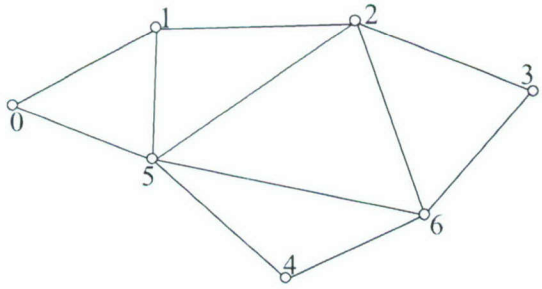


Fig. 4. *An example. Unit-disk graph G contains 9 nodes and 12 links.*

For this example, 3 steps are needed to compute the connected dominated set. They are described in detail below.

0. Initially all nodes are in state S_0.
1. Host 0 broadcasts message $< dominator(0, 0, 0) >$ and goes to state S_3. After receiving $< dominator(0, 0, 0) >$, nodes 1 and 5 broadcast messages $< dominatee(1, 0, 1) >$ and $< dominatee(5, 0, 1) >$, respectively, and go to state S_1. After receiving $< dominatee(1, 0, 1) >$ and/or $< dominatee(5, 0, 1) >$, nodes 2, 4, and 6 broadcast messages $< active(2) >$, $< active(4) >$, and $< active(6) >$, respectively, and go to state S_2. After this step, the tree contains inner node $\{0\}$ and leaves $\{1, 5\}$.
2. After the first step, W_2, W_6, W_4 contain $\{6\}$, $\{2, 4\}$, and $\{6\}$, respectively. Therefore node 2 and 4 will become dominators since they have the smallest id compared with their yellow neighbors in W_2 and W_4, respectively. Thus nodes 2 and 4 broadcast message $< dominator(2, 1, 2) >$ and $< dominator(4, 5, 2) >$, respectively, and go to state S_3. Here node 2 selects node 1 as its dominator because node 1 has a smaller id compared with node 5. After receiving $< dominator(2, 1, 2) >$, node 1 broadcasts message $< dominator(1, 0, 1) >$ and goes to state S_3; similarly node 5 broadcasts message $< dominator(5, 0, 1) >$ and goes to state S_3. After receiving $< dominator(2, 1, 2) >$ and/or $< dominator(4, 5, 2) >$ and/or $< dominator(5, 0, 1) >$, nodes 3 and 6 broadcast messages $< dominatee(3, 2, 3) >$ and $< dominatee(6, 2, 3) >$ respectively, and go to state S_1. After this step, the tree contains inner nodes $\{0, 1, 2, 4, 5\}$ and leaves $\{3, 6\}$.
3. Node 4 has no dominatee and therefor go to state S_1 to become a dominatee. The final tree contains inner nodes $\{0, 1, 2, 5\}$ and leaves $\{3, 4, 6\}$, as shown in Fig. 5.

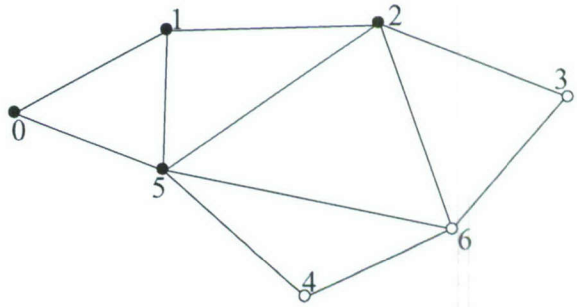


Fig. 5. *The computed connected dominating set contains nodes* $\{0, 1, 2, 5\}$. *The optimal solution contains* $\{5, 6\}$.

V. Algorithm II: Connecting a Maximal Independent Set

Algorithm I could be improved if node degree is taken into consideration. In Algorithm II, instead of growing a CDS from a leader, we compute a MIS first, then connect all nodes in the MIS. Accordingly Algorithm II contains two phases.

In Algorithm II, each node needs to maintain two more parameters: the *effective degree* d^*, which is defined to be the number of white/yellow neighbors, and the *black degree* d_b^*, which is used by a gray node to record the number of black neighbors with a higher rank. The distributed implementation of algorithm II is detailed in the following subsection.

A. Distributed Implementation

The state transition diagram for the first phase is shown in Figure 6.

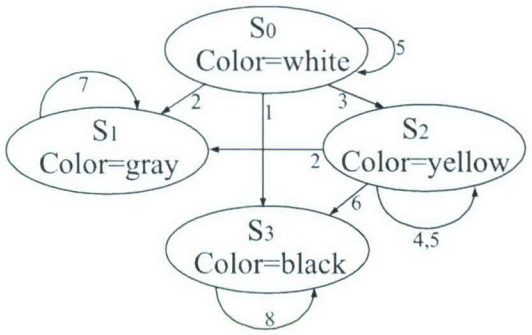


Fig. 6. *The state transition diagram for the first phase of Algorithm II for any node.*

Initially all nodes are colored white. During the execution of phase 1, each white/yellow node keeps track of its effective degree d^*, which will be updated when a neighbor changes color. To keep the degree information up-to-date, we need a new message $< degree(u, d^*) >$, by which u tells its neighbors that its effective degree is d^*. We also need the message $< blackdegree(u, d_b^*) >$, which is broadcasted by u, a dominatee. The "rank" information will be used in phase 2 to connect all members in the generated MIS. The transitions are explained below. Initially $d^* = d$, where d is the degree whose value is $|N_1[u]| - 1$.

1) u is the leader and is in state S_0. u will broadcast message $< dominator(u, u, 0) >$ to its neighbors and go to state S_3.
2) u is in state S_0 or S_2 and receives $< dominator(v, x, l) >$ from neighbor v. Here x is either the leader or $null$. u will broadcast $< dominatee(u, v, l) >$ and go to S_1.
3) u is in state S_0 and receives the message $< dominatee(v, d, l) >$. u will update d^* to $d^* - 1$, removes v from W_u if $v \in W_u$, set $rank_u$ to $l + 1$, broadcast $< active(u) >$, and then go to S_2.
4) In S_2, u will keep track of all $< dominatee(v, d, l) >$ messages broadcasted in $N_1[u]$ and remove those dominatees in W_u. u updates d^* and $rank_u$ accordingly ($rank_u = 1 + max_v\{l\}$). If there is no broadcasting in $N_1[u]$ for T_0 time unit, u will broadcast message $< degree(u, d^*) >$ if d^* is changed after last broadcasting of $< degree >$ message.
5) u is in state S_0 or S_2. If u receives message $< active(v) >$, then $W_u = W_u \cup \{v\}$; if u receives message $< degree(v, d^*) >$, u updates the local record of v's effective degree.
6) u is in state S_2 and there is no broadcasting in $N_1[u]$ for T_0 time unit. If u has biggest (d^*, id) compared with all nodes in W_u, u will broadcast $< dominator(u, null, rank_u) >$ and go to state S_3.
7) u is in state S_1. u keeps track of the number of higher rank black neighbors (number of higher rank dominators in $N_1[u]$, denoted by d_b^*). When all of its neighbors are either in S_1 or in S_3, u broadcasts the message $< blackdegree(u, d_b^*) >$.
8) u is in state S_3. u keeps track of the gray neighbor with a lower rank whose (d_b^*, id) is the biggest, in lexicographic order. This is u's candidate dominator in phase 2.

Remarks: (i) Phase 1 generates many stars. Each star consists of one black node, which serves as the center, and many gray nodes dominated by the center. Each gray node in a star sets its dom to the center. All nodes in a star have the same rank, which is also the rank of the star. The dominator of a gray node is its first black neighbor. In other words, a gray node u resides in the star centered at the black neighbor which is the first to go to state S_3. (ii) The star centered at the leader has rank 0. The rank of any other star s centered at u is one plus the highest rank of u's gray neighbors not in s. (iii) We can also understand phase 1 in the following way: phase 1 contains multiple steps. Each step generates a star. At each step except the first one, an active node (in S_2) u with maximum effective degree compared with all of its active yellow neighbors is colored black. All of u's 1-hop white/yellow neighbors (in state S_1 or S_2) are colored gray. All of u's 2-hop white neighbors (in state S_0) will go to state S_2. (iv) $< degree >$ message is used to announce the number of white/yellow neighbors of node u to all of u's white/yellow neighbors. With this information, a dominator will be elected in next step. $< blackdegree >$ message is used by gray node u to announce the number of black neighbors whose rank is higher than u to all higher rank black neighbors of u. With this information, a black node v can select a gray neighbor with a lower rank than v that connects to many stars.

After phase 1, each node is either in state S_1 or state S_3. A node in state S_1 has a dominator in S_3 while a node in S_3 does not have any dominator (except the leader). Phase 2 will designate a dominator for each black node u (in S_3). u's dominator will be the gray node v in S_1 such that $rank(v) < rank(u)$ and v has the largest (d_b^*, id) among all lower-ranked gray neighbors of u. The transition diagram for phase 2 is shown in Figure 7. We elaborate the details below.

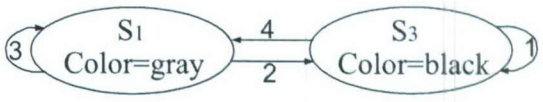


Fig. 7. *The state transition diagram for the second phase of Algorithm II for each node.*

1) After a black node u (in S_3) received $< blackdegree >$ from all of its lower-ranked gray neighbors, it broadcasts message $< dominator(u, dom_u, rank_u) >$, where dom_u is the lower-ranked gray neighbor with highest (d_b^*, id). Here $rank_u$ was set in phase 1.
2) After a gray node u (in S_1) received $< dominator(v, u, r_v) >$ from v, it will broadcast

$< dominator(u, dom_u, rank_u) >$ and go to state S_3. Here the dom_u and $rank_u$ parameters for u were set in phase 1.

3) A gray node u in S_1 will keep track of the black neighbor v with the lowest rank by listening to all $< dominator >$ broadcastings. After all black neighbors (generated in phase 1) determine their dominators, u will broadcast message $< dominatee(u, v, rank_v) >$ if v is different than its original dominator dom_u. Note that here u will be a dominatee after phase 2.
4) u is in state S_3 but no gray node selects u as the dominator, u will broadcast message $< dominatee(u, dom_u, rank_u) >$ and go to state S_1.

Remarks: (i) Phase 2 assigns a dominator v to the center u of each star generated in phase 1 except the star centered at the leader. v is a gray node satisfying the following conditions: v is located in a star with a rank lower than u and v is adjacent to maximum number of stars with a higher rank than v. (ii) After phase 2, the dominator of a gray node u is always the black neighbor with the lowest rank. (iii) Phase 2 is a local process which happens among adjacent stars with different ranks. After a black node u received $< blackdegree >$ from all of its lower-ranked gray neighbors, it can select its dominator and broadcast message $< dominator(u, dom_u, rank_u) >$. This means a black node can initiate phase 2 immediately after all necessary information is available. There is no explicit start time for phase 2. Phase 2 is used to connect all stars.

Remarks: (i) There are many similarities between Algorithm I and II. Both algorithms grow a maximal independent set (MIS) from the leader (see Theorem 4.1 and Lemma 5.1). (ii) The major differences are: (a) The criteria to decide whether an active node can be elected to be a member of the MIS is different. In Algorithm I, an active node with a smallest id compared with all of its active neighbors will be added into the MIS. In Algorithm II, an active node with maximum effective degree compared with all of its active neighbors will be added into the MIS. (b) The time when the dominators of the nodes in the MIS are determined is different. Algorithm I assigns a dominator to each node in the MIS when the node satisfies the condition in (a); Algorithm II delays the determination of the dominator to phase 2. Thus Algorithm II can pick the gray node that connects to a larger number of elements in MIS.

B. An example

Now we show how to apply Algorithm II to the example in subsection IV-B. Initially all nodes are in state S_0. Phase 1 contains 2 steps:

1) Host 0 broadcasts message $< dominator(0, 0, 0) >$ and goes to state S_3. After receiving $< dominator(0, 0, 0) >$, nodes 1 and 5 broadcast messages $< dominatee(1, 0, 0) >$ and $< dominatee(5, 0, 0) >$, respectively, and go to state S_1. After receiving $< dominatee(1, 0, 0) >$ and/or $< dominatee(5, 0, 0) >$, nodes 2, 4, and 6 broadcast messages $< active(2) >$, $< active(4) >$, and $< active(6) >$, respectively, and go to state S_2. After all these broadcastings, nodes 2, 4, and 6 broadcast messages $< degree(2, 2) >$, $< degree(4, 1) >$, and $< degree(6, 3) >$, respectively. This step generates the star $(0, \{1, 5\})$.
2) Nodes 2, 4, 6 maintain the list of active neighbors and the corresponding effective degrees, which are $\{(3, 6)\}$, $\{(2, 2), (1, 4)\}$, and $\{(3, 6)\}$, respectively. The winner of the competition is node 6 since it has the highest (d^*, id), thus node 6 broadcasts message $< dominator(6, null, 1) >$ and go to state S_3. After receiving $< dominator(6, null, 1) >$, nodes 2, 3, and 4 broadcast messages $< dominatee(2, 6, 1) >$, $< dominatee(3, 6, 1) >$, and $< dominatee(4, 6, 1 >$, respectively, and go to state S_1. Then all nodes 1, 2, 3, 4, and 5 broadcast messages $< blackdegree(5, 1) >$, $< blackdegree(1, 1) >$, $< blackdegree(2, 0) >$, $< blackdegree(3, 0) >$, and $< blackdegree(4, 0) >$, respectively. This step generates the star $(6, \{2, 3, 4\})$.

After phase 1, nodes $\{0, 6\}$ are in state S_3 and all other nodes are in state S_1. Phase 2 contains the following step:

3. After receiving $< blackdegree >$ message from node 5, the only lower-ranked gray neighbor, node 6 broadcasts message $< dominator(6, 5, 1) >$. After receiving $< dominator(6, 5, 1) >$, node 5 broadcasts message $< dominator(5, 0, 0) >$ and goes to state S_3. After receiving $< dominator(5, 0, 0) >$, nodes 2 and 4 will selects node 5 as its dominator. The algorithm terminates. The resultant CDS contains nodes $\{0, 5, 6\}$.

C. Performance analysis

In this subsection, we study the performance of Algorithm II.

Lemma 5.1: After phase 1, all nodes in state S_3 form a MIS. In other words, the centers of all stars form a MIS.

Proof. A node is colored black only from white. No two white neighboring nodes could be colored black at the same time since they must have different (d^*, id). Whenever a node is colored black, all of its neighbors are colored gray. Once a node is colored gray, it remains in color gray during Phase 1. ■

Lemma 5.2: Phase 2 generates a connected dominating set.

Proof. Note that phase 1 generates many stars whose centers form a MIS. Two stars are adjacent if a gray node in one star is a neighbor of the center of the other star. Since the input graph is connected, a star except the lowest-ranked one (the one with the leader) centered at the leader has at least one adjacent star. Each star has a rank which is higher than the rank of its adjacent stars. Phase 2 connects a star to a lower-ranked adjacent star by turning a gray vertex in the lower-ranked star to black. Thus after phase 2, each black vertex (the center of a star) has a path consisting of only black nodes to the leader. Therefore all these black vertices form a connected dominating set. ■

Lemma 5.3: In phase 2, the gray node with maximum (d_b^*, id) will connect to d_b^* number of lower-ranked stars.

Proof. Let u be the gray node with maximum (d_b^*, id). Phase 1 ensures that u belongs to the star whose center is the first black neighbor of u. For all other black neighbors (centers of some stars) of u, their ranks must be greater than that of u since a black node in phase 1 always assigns its rank to be one plus the rank of its highest rank dominatee neighbor. Thus all these d_b^* nodes will select u as their dominators in phase 2. In other words, u connects to d_b^* number of lower-ranked stars. ■

Theorem 5.1: The connected dominating set generated in Algorithm II has a size at most $(14n_k + 8)opt + 4n_k - 1$, where opt is the size of any optimal MCDS for the given instance.

Proof. From Lemma 5.1, phase 1 computes a MIS. Let A be this MIS with a size $|A|$. From Theorem 3.1, $|A| \leq (7n_k + 4)opt + 2n_k + 1$. Note that in phase 2, at most $|A| - 1$ nodes in state S_1 will go to state S_3. Now we consider two cases here.

First, if there exist a gray vertice with $d_b^* \geq 2$ at the beginning of phase 2, from Lemma 5.3, the gray vertex u with maximum (d_b^*, id) will connect d_b^* stars to the higher rank star u resides in phase 2. Therefor the number of nodes changing state from S_1 to S_3 in phase 2 is at most $|A| - 2$. Thus the total number of nodes in state S_3 is at most $|A| + |A| - 2 \leq (7n_k + 4)opt + 2n_k + 1 + (7n_k + 4)opt + 2n_k + 1 - 2 \leq (14n_k + 8)opt + 4n_k - 1$.

Secondly, if all gray vertices have $d_b^* \leq 1$ at the beginning of phase 2, then the number of nodes changing state from S_1 to S_3 in phase 2 is exactly $|A| - 1$. Since the dominator of any gray vertex u is its first black neighbor, and all other black neighbors of u have higher rank than u, thus the total number of black neighbors u has is at most 2. In other words, any node in an optimal MCDS is either in A or adjacent to at most 2 vertices in A and any vertex in A is dominated by a vertex in the MCDS. Therefor in this case, $|A| \leq 2 \cdot opt$. Thus the total number of black nodes will be $2 \cdot opt + 2 \cdot opt - 1 < 4 \cdot opt$. ■

Theorem 5.2: Algorithm II has message complexity $O(n\Delta)$, where n is the total number of vertices and Δ is the maximum node degree.

Proof. In Algorithm II, the message complexity is dominated by the $< degree >$ messages broadcasted by white vertices in S_2 in phase 1 since each node broadcasts each of the other messages at most twice. Therefore the message complexity is $O(n\Delta)$. ■

VI. Conclusion

In this paper, we studied the problem of constructing minimum connected dominating set in geometric k-disk graphs. We first derived the upper bound of the size of any maximal independent set compared to that of a MCDS. Then we proposed two distributed approximation algorithms and studied their performance theoretically.

As a future research, we intend to improve the upper bound and design better approximation algorithms for MCDS in geometric k-disk graphs. Note that a polynomial time approximation scheme (PTAS) exists for unit-disk graphs [15] but whether a PTAS exists or not for the general geometric k-disk graphs is still open. This is another problem we intend to target in the future.

References

[1] M. D. J. Blum and X. Cheng, *Handbook of Combinatorial Optimization*. Kluwer Academic Publisher, 2004, ch. Applications of Connected Dominating Sets in Wireless Networks, pp. 329–369.

[2] X. Cheng, M. Ding, D. H. Du, and X. Jia, "Virtual backbone construction in multihop ad hoc wireless networks," *Wireless Communications and Mobile Computing*, vol. 6, pp. 183–190, 2006.

[3] B. N. Clark, C. J. Colbourn, and D. S. Johnson, "Unit disk graphs," *Discrete Mathematics*, vol. 86, pp. 165–177, 1990.

[4] M. T. Thai, F. Wang, D. Liu, S. Zhu, and D.-Z. Du, "Connected dominating sets in wireless networks with different transmission ranges," *IEEE Transactions on Mobile Computing*, vol. 6, no. 7, pp. 721–730, 2007.

[5] M. R. Garey and D. S. Johnson, *Computers and Intractability: A Guide to teh Theory of NP-Completeness*. W. H. Freeman, 1979.

[6] S. Guha and S. Khuller, "Approximation algorithms for connected dominating sets," *Algorithmica*, vol. 20, no. 4, pp. 374–387, 1998.

[7] L. Ruan, H. Du, X. Jia, W. Wu, Y. Li, and K.-I. Ko, "A greedy approximation for minimum connected dominating sets," *Theor. Comput. Sci.*, vol. 329, no. 1-3, pp. 325–330, 2004.

[8] J. Wu and H. Li, "On calculating connected dominating set for efficient routing in ad hoc wireless networks," in *DIALM '99: Proceedings of the 3rd international workshop on Discrete algorithms and methods for mobile computing and communications*. New York, NY, USA: ACM, 1999, pp. 7–14.

[9] P.-J. Wan, K. M. Alzoubi, and O. Frieder, "Distributed construction of connected dominating set in wireless ad hoc networks," in *Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2002)*, 2002, pp. 1597–1604.

[10] W. Wu, H. Du, X. Jia, Y. Li, and S. C.-H. Huang, "Minimum connected dominating sets and maximal independent sets in unit disk graphs," *Theor. Comput. Sci.*, vol. 352, no. 1, pp. 1–7, 2006.

[11] M. Cardei, M. X. Cheng, X. Cheng, and D.-Z. Du, "Connected domination in ad hoc wireless networks," in *International Conference on Computer Science and Informatics (CS&I 2002)*, 2002, pp. 251–255.

[12] K. Alzoubi, P.-J. Wan, and O. Frieder, "New distributed algorithm for connected dominating set in wireless ad hoc networks," in *HICSS '02: Proceedings of the 35th Annual Hawaii International Conference on System Sciences (HICSS'02)-Volume 9*. Washington, DC, USA: IEEE Computer Society, 2002, p. 297.

[13] Y. Li, M. T. Thai, F. Wang, C.-W. Yi, P.-J. Wan, and D.-Z. Du, "On greedy construction of connected dominating sets in wireless networks: Research articles," *Wirel. Commun. Mob. Comput.*, vol. 5, no. 8, pp. 927–932, 2005.

[14] K. M. Alzoubi, P.-J. Wan, and O. Frieder, "Message-optimal connected dominating sets in mobile ad hoc networks," in *MobiHoc '02: Proceedings of the 3rd ACM international symposium on Mobile ad hoc networking & computing*. New York, NY, USA: ACM, 2002, pp. 157–164.

[15] X. Cheng, X. Huang, D. Li, W. Wu, and D.-Z. Du, "A polynomial-time approximation scheme for the minimum-connected dominating set in ad hoc wireless networks," *Networks*, vol. 42, no. 4, pp. 202–208, 2003.

[16] I. Cidon and O. Mokryn, "Propagation and leader election in a multihop broadcast environment," in *DISC '98: Proceedings of the 12th International Symposium on Distributed Computing*. London, UK: Springer-Verlag, 1998, pp. 104–118.

[17] L. Subramanian, R. H. Katz, V. Roth, S. Shenker, and I. Stoica, "Reliable broadcast in unknown fixed-identity networks," EECS Department, University of California, Berkeley, Tech. Rep. UCB/CSD-04-1358, 2004. [Online]. Available: http://www.eecs.berkeley.edu/Pubs/TechRpts/2004/6505.html

[18] C.-Y. Koo, V. Bhandari, J. Katz, and N. H. Vaidya, "Reliable broadcast in radio networks: the bounded collision case," in *PODC '06: Proceedings of the twenty-fifth annual ACM symposium on Principles of distributed computing*. New York, NY, USA: ACM, 2006, pp. 258–264.